



THE UNITED STATES PATENT AND TRADEMARK OFFICE

Attorney Docket No. 040301/0575

Applicant: Shigeo MATSUZAWA et al.

Title: ROUTER DEVICE AND CUT-THROUGH PATH CONTROL
METHOD FOR REALIZING LOAD BALANCING AT
INTERMEDIATE ROUTERS

Serial No.: 09/429,632

Filed: October 29, 1999

Examiner: Ho, Chuong T.

Art Unit: 2664

RECEIVED

MAY 21 2003

Technology Center 2600

CLAIM FOR CONVENTION PRIORITY

Commissioner for Patents
PO Box 1450
Alexandria, Virginia 22313-1450

Sir:

The benefit of the filing date of the following prior foreign application filed in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of said original foreign application:

JAPAN Patent Application No. 10-311315 filed October 30, 1998.

Respectfully submitted,

Date

19 May, 2003

Phillip J. Articola

Phillip J. Articola
Registration No. 38,819

FOLEY & LARDNER
Washington Harbour
3000 K Street, N.W., Suite 500
Washington, D.C. 20007-5143
Telephone: (202) 672-5300
Facsimile: (202) 672-5399

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
in this Office.

出 願 年 月 日

Date of Application:

1998年10月30日

出 願 番 号

Application Number:

平成10年特許願第311315号

出 願 人

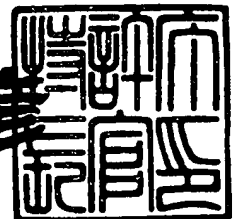
Applicant(s):

株式会社東芝

1999年10月15日

特許庁長官
Commissioner,
Patent Office

近 藤 隆 彦



CERTIFIED COPY OF
PRIORITY DOCUMENT 出証番号 出証特平11-3069858

【書類名】 特許願

【整理番号】 A009806895

【提出日】 平成10年10月30日

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/00

【発明の名称】 カットスループス制御方法及びルータ装置並びに通信方法

【請求項の数】 21

【発明者】

【住所又は居所】 東京都日野市旭が丘3丁目1番地の1 株式会社東芝日野工場内

【氏名】 松澤 茂雄

【発明者】

【住所又は居所】 東京都日野市旭が丘3丁目1番地の1 株式会社東芝日野工場内

【氏名】 今野 徹

【発明者】

【住所又は居所】 東京都日野市旭が丘3丁目1番地の1 株式会社東芝日野工場内

【氏名】 茂木 章善

【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

【代理人】

【識別番号】 100058479

【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1
【物件名】 要約書 1
【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 カットスループス制御方法及びルータ装置並びに通信方法

【特許請求の範囲】

【請求項 1】

マルチパスが存在するカットスループスの設定に際し、自装置が関与するカットスループスの設定状況に関する全部または所定の一部の情報に基づいて、次段ルータとし得る複数のルータ装置のうちから負荷分散に寄与するようにルータ装置を選択し、

この選択されたルータ装置を次段ルータとするカットスループスの設定のための所定の制御を行うことを特徴とするカットスループス制御方法。

【請求項 2】

ある宛先のためのカットスループスを設定する際に、次段ルータとなり得るルータ装置が複数存在する場合、宛先を同じくする既設定のカットスループス数に基づいて、該次段ルータとなり得る各ルータ装置へのカットスループス数が均等的に配分されるように、該設定すべきカットスループスにおいて次段ルータとするルータ装置を選択することを特徴とするカットスループス制御方法。

【請求項 3】

ある整数値を、前記宛先のためのカットスループスにおける次段ルータとなり得るルータ装置の数で除したときに出現し得る、0 から始まる各剰余の値夫々を、該次段ルータとなり得るルータ装置に 1 つずつ割り当てておき、

前記選択の際、前記既設定のカットスループス数を前記ルータ装置の数で除して得られた剰余の値に対して割り当てられたルータ装置を選択することを特徴とする請求項 2 に記載のカットスループス制御方法。

【請求項 4】

ある宛先のためのカットスループスを設定する際に、次段ルータとなり得るルータ装置が複数存在する場合、宛先を同じくする既設定のカットスループス数に基づいて、該次段ルータとなり得る各ルータ装置へのカットスループス数が各ルータ装置に対するリンク速度に応じて公平に配分されるように、該設定すべきカットスループスにおいて次段ルータとする次段ルータとするルータ装置を選択す

ることを特徴とするカットスループス制御方法。

【請求項 5】

ある整数値を、前記宛先のためのカットスループスにおける次段ルータとなり得る各ルータ装置に対するリンク速度の比を表すもしくは近似する整数比の各要素の総和で除したときに出現し得る、0 から始まる各剰余の値夫々を、該次段ルータとなり得るルータ装置に、前記リンク速度に応じた個数ずつ割り当てておき

、
前記選択の際、前記既設定のカットスループス数を前記整数比の各要素の総和で除して得られた剰余の値に対して割り当てられたルータ装置を選択することを特徴とする請求項 4 に記載のカットスループス制御方法。、

【請求項 6】

前記選択されたルータ装置に前記カットスループスを設定するためのメッセージを送出し、

前記カットスループスが設定されたならば、該カットスループスを利用するための情報設定を行うことを特徴とする請求項 1 ないし 6 のいずれか 1 項に記載のカットスループス制御方法。

【請求項 7】

前記選択されたルータ装置への他の既設定のカットスループスが未だ存在しない場合には、前記選択したルータ装置に前記カットスループスを設定するためのメッセージを送出し、該カットスループスが設定されたならば、該カットスループスを利用するための情報設定を行い、

前記選択されたルータ装置への他の既設定のカットスループスが既に存在する場合には、該既に設定されているカットスループスに該設定すべきカットスループスをマージするための情報設定を行うことを特徴とする請求項 1 ないし 6 のいずれか 1 項に記載のカットスループス制御方法。

【請求項 8】

前記カットスループスの設定は、上流側のノード装置からカットスループスを設定するためのメッセージを受信したことを契機として開始することを特徴とする請求項 1 ないし 7 のいずれか 1 項に記載のカットスループス制御方法。

【請求項9】

所定のタイミングで、自装置からのカットスループスの次段ノードとなっている各ノード装置へのトラフィック量を調べ、

前記各ノード装置へのトラフィック量の不均衡状態が、予め規定された許容範囲を越えている場合には、トラフィック量のより少ない方のノード装置を次段ノードとしているマルチパスの存在するカットスループスであって該トラフィック量に関する所定の条件を満足するものを選択し、

この選択したカットスループスの次段ノードを、トラフィック量の多い方のノード装置に設定変更することを特徴とする請求項1ないし8のいずれか1項に記載のカットスループス制御方法。

【請求項10】

前記選択および設定変更を、前記不均衡状態が前記許容範囲内に収まるまで、または前記選択ができなくなるまで、繰り返し行うことを特徴とする請求項9に記載のカットスループス制御方法。

【請求項11】

前記選択において、最適な1または複数のカットスループスの解を求めることを特徴とする請求項9に記載のカットスループス制御方法。

【請求項12】

自装置が関与するカットスループスのトラフィック状況に関する全部または所定の一部の情報に基づいて、自装置において経路変更可能なカットスループスのうち経路変更した場合に負荷分散に寄与するものを選択し、

この選択されたカットスループスの経路を、負荷分散に寄与するように経路変更することを特徴とするカットスループス制御方法。

【請求項13】

マルチパスが存在するカットスループスを設定するに際し、自装置が関与するカットスループスの設定状況に関する全部または所定の一部の情報に基づいて、次段ルータとし得る複数のルータ装置のうちから負荷分散に寄与するようにルータ装置を選択する手段と、

この選択されたルータ装置を次段ルータとするカットスループスの設定のため

の所定の制御を行う手段とを具備したことを特徴とするルータ装置。

【請求項 14】

自装置が関与するカットスループスのトラフィック状況に関する全部または所定の一部の情報に基づいて、自装置において経路変更可能なカットスループスのうち経路変更した場合に負荷分散に寄与するものを選択する手段と、

この選択されたカットスループスの経路を、負荷分散に寄与するように経路変更する手段とを具備したことを特徴とするルータ装置。

【請求項 15】

プライベートアドレス空間から成るプライベートネットワークに接続されパケットを転送する手段を有するプライベート処理モジュールと、非プライベートアドレス空間から成る非プライベートネットワークに接続されパケットをラベルスイッチする手段を有する非プライベート処理モジュールと、を備えるシステムにおける通信方法であって、

前記非プライベート処理モジュールは、前記プライベート処理モジュールが接続されたプライベートネットワークと前記プライベートアドレス空間を共有する他のプライベートネットワークを収容する他のシステムが前記非プライベートネットワークに存在することを認識した場合に、当該他のシステムへ向けて前記非プライベートネットワークを経由するラベルスイッチパスを設定するためのプロトコル通信を行い、

前記非プライベート処理モジュールは、設定された前記ラベルスイッチパスを前記プライベート処理モジュールまで延長し、前記プライベート処理モジュールに前記他のシステムのプライベートアドレスを通知し、

前記プライベート処理モジュールは、前記他のシステムのプライベートアドレス宛のパケットを、延長された前記ラベルスイッチパスへ転送することを特徴とする通信方法。

【請求項 16】

前記プライベート処理モジュールと前記非プライベート処理モジュールとの間に予め複数のパスを用意しておき、前記ラベルスイッチパスの延長は、この複数のパスから選択した一つのパスと、設定された前記ラベルスイッチングパスとを

、前記非プライベート処理モジュール内で対応付けることにより行い、前記他のシステムのプライベートアドレスは、選択された前記一つのパスと対応付けて通知されることを特徴とする請求項 15 に記載の通信方法。

【請求項 17】

前記ラベルスイッチパスの延長は、前記ラベルスイッチパスが前記他のシステムに到達するように設定されたことが確認された後に行うことを特徴とする請求項 15 に記載の通信方法。

【請求項 18】

前記他のシステムのプライベートアドレスの通知は、パケットを前記他のシステムへ運ぶための前記ラベルスイッチと、パケットを前記他のシステムから運ぶための別のラベルスイッチとの双方が設定された場合に行うことを特徴とする請求項 15 に記載の通信方法。

【請求項 19】

前記システムは、前記プライベートアドレス空間とは異なるプライベートアドレス空間から成る第二のプライベートネットワークに接続されパケットを転送する手段を有する第二のプライベート処理モジュールをも備えるものであり、

前記非プライベート処理モジュールは、前記第二のプライベート処理モジュールが接続された第二のプライベートネットワークと前記異なるプライベートアドレス空間を共有する他の第二のプライベートネットワークを収容する他のシステムが前記非プライベートネットワークに存在することを認識した場合にも、当該他のシステムへ向けて前記非プライベートネットワークを経由するラベルスイッチパスを設定するためのプロトコル通信を行い、

前記非プライベート処理モジュールは、設定された前記ラベルスイッチパスを前記第二のプライベート処理モジュールまで延長し、前記第二のプライベート処理モジュールに前記他のシステムのプライベートアドレスを通知し、

前記第二のプライベート処理モジュールは、前記他のシステムのプライベートアドレス宛のパケットを、延長された前記ラベルスイッチパスへ転送することを特徴とする請求項 15 に記載の通信方法。

【請求項 20】

ラベルスイッチパスをその内部に設定可能なラベルスイッチネットワークと、それ以外の非ラベルスイッチネットワークとを含み、プライベートアドレス空間を共有する少なくとも二つのプライベートネットワークの間に位置する非プライベートネットワークにおいて、仮想的なプライベートネットワークを構成するための通信方法であって、

前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータが、前記プライベートネットワークの一方から当該ルータへのラベルスイッチパスを設定するのに用いられた制御メッセージの内容に基づき、該ラベルスイッチパスから受信したパケットの宛先プライベートアドレスが前記プライベートネットワークの他方であることを認識し、受信した該パケットを前記プライベートネットワークの他方に対応する非プライベートアドレスでカプセル化して送信することを特徴とする通信方法。

【請求項 21】

前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータが、前記プライベートネットワークの一方から当該ルータへのラベルスイッチパスを設定するのに用いられた制御メッセージの内容に基づき、自身が前記プライベートネットワークの境界ルータであると前記非プライベートネットワーク内に広告し、

前記プライベートネットワークの他方を収容する前記非プライベートネットワークに接続されたルータは、この広告に基づいて前記プライベートネットワークの他方からのパケットを前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータの非プライベートアドレスでカプセル化して送信し、

このカプセル化されたパケットを受信した前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータは、該カプセル化を解いて該パケットを当該ルータから前記プライベートネットワークの一方へのラベルスイッチパスへ送信することを特徴とする請求項 20 に記載の通信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、カットスループスを設定するルータ装置及びカットスループス制御方法並びにラベルスイッチ機能を有するルータ装置の通信方法に関する。

【0002】

【従来の技術】

IP（インターネットプロトコル）ヘッダを解析することによりデータグラムを転送するルータ装置では、次の転送先を決めるためにルーチングプロトコルを使用している。その一つとして使用されているOSPF（Open Shortest Path Fast；例えば、J. Moy, "OSPF Version 2", Internet RFC 2328, April 1998参照）等では、最終宛先ネットワークまたはホストまでの同一のコストの経路について、次の転送先が複数存在する場合には、次の転送先情報を複数保持することが可能である。ここで、コストとは例えば途中経由するルータ数などの情報であり、コストが低い場合に最短経路で到着することを意味する。

【0003】

この機能は負荷分散（ロードバランス）などで使用し、例えば、転送するルータ装置で、複数の次の転送先ルータ情報を持つ場合には、それぞれについて均一に出力することで、片方のルータ装置だけが高負荷にならないようにしている。ここで、注意すべき点は、ロードバランス時には、2ホスト間のデータ転送において、データグラムの順序が逆転しないようにすることが必要である。

【0004】

一方、IP等のデータグラム転送を高速化するための一手法として考案されているラベルスイッチの技術では、データグラム転送中にはIPヘッダ情報を見ないため、2ホスト間のデータ転送におけるデータグラムの順序保証は困難である。

【0005】

そこで、現在はIPヘッダによる転送を行う装置（ここでは、エッジルータと呼ぶ）から、複数のカットスルー転送を行うためのパス（ラベルスイッチパスと

呼ぶ)を設定し、エッジルータで複数設定したカットスループスのいずれかにデータグラムを転送するという手法が取られる。

【0006】

この手法におけるラベルスイッチパス設定としては、エッジルータから設定を開始するときに、各パスを強制的に通るような設定(Explicit Route)が存在する。

【0007】

Explicit Routeによる設定の場合、途中のルータ装置での経路変更による、指定した経路(ルータ装置)が削除された場合等において、必ずエッジルータが認識する必要がある、指定された経路が削除されたことを検出したルータからエッジルータに対し通知する際に、ルーチングプロトコルもしくはカットスルーを生成するためのプロトコル情報を使用する等、特別な手段が必要とされる場合がある。

【0008】

また、従来の方法では、自身には複数の経路を持たない場合でも、複数経路を認識する必要がある、ルーチングプロトコルによっては困難な場合があり、ネットワーク管理者がネットワーク設計時に登録しなければならない可能性がある。

【0009】

【発明が解決しようとする課題】

以上のように従来、カットスループスに関するロードバランスを実行する場合、ロードバランスのために経路変更する該当カットスループスのエッジルータがこれに関与する必要があるという問題点があった。また、このために、制御や実装が複雑になる、効果的なロードバランスの実現が難しい、などの不具合があった。

【0010】

本発明は、上記事情を考慮してなされたもので、ロードバランスを実行する場合に、実際にマルチパス情報を持つルータ装置にて行い、エッジルータにて特別な処理が必要とならないようにしたカットスループス制御方法及びルータ装置を提供することを目的とする。

【0011】

【課題を解決するための手段】

本発明（請求項1）に係るカットスループス制御方法は、マルチパスが存在するカットスループスの設定に際し、自装置が関与するカットスループスの設定状況に関する全部または所定の一部の情報（例えば、設定しようとしているカットスループスと宛先を同じくする既設定のカットスループス数、選択し得る各ルータ装置に対するリンク速度）に基づいて、次段ルータとし得る複数のルータ装置のうちから負荷分散に寄与するようにルータ装置を選択し、この選択されたルータ装置を次段ルータとするカットスループスの設定のための所定の制御を行うことを特徴とする。

【0012】

ここで、ルータ装置とは、IPなどのネットワーク層の経路表を持つもののことを言うものとする。

【0013】

なお、負荷分散に寄与するようにルータ装置を選択する手順としては、その時点で最も負荷分散に寄与するルータ装置があるならばそれを選択するような手順だけでなく、一定のスパンで見ても負荷分散させるような手順をも意味する。簡単な例を示すと、例えば、次段ルータとし得る3つのルータ装置（ルータ1、ルータ2、ルータ3）について、ルータ1、2、3、1、2、3、…の順番に選択することになるような手順が前者に該当し、ルータ1、1、2、2、3、3、1、1、2、2、3、3、…の順番に選択することになるような手順が後者に該当する。

【0014】

また、負荷分散は、宛先を同じくするマルチパスで設定されたカットスループスの範囲毎に行っても良いし、全カットスループスに渡って行っても良いし、また、その他にも種々の方法が考えられる。

【0015】

また、種々のファクターを考慮してより実効的な負荷分散を実現するようにしてもよい。

【0016】

本発明（請求項2）に係るカットスループス制御方法は、ある宛先のためのカットスループスを設定する際に、次段ルータとなり得るルータ装置が複数存在する場合、宛先を同じくする（この場合、次段ルータとなり得る複数のルータ装置が同じになる）既設定のカットスループス数に基づいて、該次段ルータとなり得る各ルータ装置へのカットスループス数が均等的に配分されるように、該設定すべきカットスループスにおいて次段ルータとするルータ装置を選択することを特徴とする。

【0017】

なお、前述と同様、均等的に配分されるようにルータ装置を選択する手順としては、その時点で最も均等化に寄与するルータ装置があるならばそれを選択するような手順だけでなく、一定のスパンで見て均等化させるような手順をも意味する。

【0018】

また、好ましくは、ある整数値を、前記宛先のためのカットスループスにおける次段ルータとなり得るルータ装置の数で除したときに出現し得る、0から始まる各剰余の値夫々を、該次段ルータとなり得るルータ装置に1つずつ割り当てておき、前記選択の際、前記既設定のカットスループス数を前記ルータ装置の数で除して得られた剰余の値に対して割り当てられたルータ装置を選択するようにしてもよい。簡単な例を示すと、例えば、次段ルータとし得る2つのルータ装置（ルータ1、ルータ2）について、剰余が0の場合にはルータ1を選択し、剰余が1の場合にはルータ2を選択するものとする、ルータ1、2、1、2、…の順番に選択することになる。

【0019】

本発明（請求項4）に係るカットスループス制御方法は、ある宛先のためのカットスループスを設定する際に、次段ルータとなり得るルータ装置が複数存在する場合、宛先を同じくする（この場合、次段ルータとなり得る複数のルータ装置が同じになる）既設定のカットスループス数に基づいて、該次段ルータとなり得る各ルータ装置へのカットスループス数が各ルータ装置に対するリンク速度に応

じて公平に配分されるように、該設定すべきカットスループスにおいて次段ルータとする次段ルータとするルータ装置を選択することを特徴とする。

【0020】

なお、前述と同様、各ルータ装置に対するリンク速度に応じて公平に配分されるようにルータ装置を選択する手順としては、その時点で最も公平化に寄与するルータ装置があるならばそれを選択するような手順だけでなく、一定のスパンで見ても公平化させるような手順をも意味する。

【0021】

また、好ましくは、ある整数値を、前記宛先のためのカットスループスにおける次段ルータとなり得る各ルータ装置に対するリンク速度の比を表すもしくは近似する整数比の各要素の総和で除したときに出現し得る、0から始まる各剰余の値夫々を、該次段ルータとなり得るルータ装置に、前記リンク速度に応じた個数ずつ割り当てておき、前記選択の際、前記既設定のカットスループス数を前記整数比の各要素の総和で除して得られた剰余の値に対して割り当てられたルータ装置を選択するようにしてもよい。簡単な例を示すと、例えば、次段ルータとし得る3つのルータ装置（ルータ1、ルータ2、ルータ3）について、ルータ1に対するリンク速度：ルータ2に対するリンク速度：ルータ3に対するリンク速度＝1：2：3とすると、剰余が0、3または5の場合にはルータ3を選択し、剰余が1または4の場合にはルータ2を選択し、剰余が2の場合にはルータ1を選択するものとする、ルータ3、2、1、3、2、3、3、2、1、3、2、3、…の順番に選択することになる。また、剰余が0、1または2の場合にはルータ3を選択し、剰余が3または4の場合にはルータ2を選択し、剰余が5の場合にはルータ1を選択するものとして、ルータ3、3、3、2、2、1、3、3、3、2、2、1、…の順番に選択するようにすることもできる。

【0022】

さらに、以上の各構成において、好ましくは、前記選択されたルータ装置に前記カットスループスを設定するためのメッセージを送出し、前記カットスループスが設定されたならば、該カットスループスを利用するための情報設定を行うようにしてもよい。これは、マージを行わない場合である。

【0023】

あるいは、好ましくは、前記選択されたルータ装置への他の既設定のカットスループスが未だ存在しない場合には、前記選択したルータ装置に前記カットスループスを設定するためのメッセージを送出し、該カットスループスが設定されたならば、該カットスループスを利用するための情報設定を行い、前記選択されたルータ装置への他の既設定のカットスループスが既に存在する場合には、該既に設定されているカットスループスに該設定すべきカットスループスをマージするための情報設定を行うようにしてもよい。これは、マージを行う場合である。

【0024】

以上において、好ましくは、前記カットスループスの設定は、上流側のノード装置からカットスループスを設定するためのメッセージを受信したことを契機として開始するようにしてもよい。

【0025】

また、以上において、好ましくは、所定のタイミング（例えば、一定間隔で周期的に、何らかのイベントが発生したことを契機として、など）で、自装置からのカットスループスの次段ノードとなっている各ノード装置へのトラフィック量を調べ、前記各ノード装置へのトラフィック量の不均衡状態が、予め規定された許容範囲を越えている場合には、トラフィック量のより少ない方のノード装置を次段ノードとしているマルチパスの存在するカットスループスであって該トラフィック量に関する所定の条件を満足するもの（例えば、そのカットスループスの経路を設定変更したことによってトラフィック量の多いノード装置と少ないノード装置の関係を逆転させないもの、そのカットスループスの経路を設定変更したことによって前記不均衡状態が前記許容範囲内に収まることになるようなもの、など）を選択し、この選択したカットスループスの次段ノードを、トラフィック量の多い方のノード装置に設定変更するようにしてもよい。

【0026】

この場合に、好ましくは、前記選択および設定変更を、前記不均衡状態が前記許容範囲内に収まるまで、または前記選択ができなくなるまで、繰り返し行うようにしてもよい。

【0027】

あるいは、好ましくは、前記選択において、最適な1または複数のカットスループスの解を求めるようにしてもよい。

【0028】

本発明（請求項12）に係るカットスループス制御方法は、自装置が関与するカットスループスのトラフィック状況に関する全部または所定の一部の情報に基づいて、自装置において経路変更可能なカットスループスのうち経路変更した場合に負荷分散に寄与するものを選択し、この選択されたカットスループスの経路を、負荷分散に寄与するように経路変更するようにしてもよい。

【0029】

本発明（請求項13）に係るルータ装置は、マルチパスが存在するカットスループスを設定するに際し、自装置が関与するカットスループスの設定状況に関する全部または所定の一部の情報に基づいて、次段ルータとし得る複数のルータ装置のうちから負荷分散に寄与するようにルータ装置を選択する手段と、この選択されたルータ装置を次段ルータとするカットスループスの設定のための所定の制御を行う手段とを具備したことを特徴とする。

【0030】

本発明（請求項14）に係るルータ装置は、自装置が関与するカットスループスのトラフィック状況に関する全部または所定の一部の情報に基づいて、自装置において経路変更可能なカットスループスのうち経路変更した場合に負荷分散に寄与するものを選択する手段と、この選択されたカットスループスの経路を、負荷分散に寄与するように経路変更する手段とを具備したことを特徴とする。

【0031】

本発明によれば、ロードバランスを実行する場合に、エッジルータの関与なしに、実際にマルチパス情報を持つルータ装置のみにて行うことができる。

【0032】

また、制御や実装の複雑化を回避し、効果的なロードバランスを実現することができる。

【0033】

なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【0034】

また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【0035】

【発明の実施の形態】

以下、図面を参照しながら発明の実施の形態を説明する。

【0036】

図1は、ラベルスイッチを導入したネットワークであってルータ装置が次段情報を複数持ち得るようなネットワークの一例を示すものである。

【0037】

ルータ装置101～103は、各ルータ間にて、転送するデータグラムフローの情報と下位レイヤの情報について共通の認識を得ることで、IPヘッダ情報を見ることなく、下位レイヤ情報だけに基づいて高速にデータグラム転送を行うことを可能とするパス（以下、カットスルーパスと呼ぶ）を生成する機能を有する装置である。

【0038】

特定のデータグラムの情報と下位レイヤ情報（以下、ラベルと呼ぶ）について共通の認識を得る手段としては、例えば、FANP (Flow Attribute Notification Protocol)、TDP (Tag Distribution Protocol)、LDP (Label Distribution Protocol) 等のプロトコルが用いられる。

【0039】

ルータ装置104は、ルータ装置102、103やその先のネットワーク120と接続される装置であり、ルータ装置102およびルータ装置103との間で

FANP、TDP、LDP等のプロトコルが動作しており、カットスループスの設定やカットスループスの最終段としての機能を有する。

【0040】

ルータ装置111～11Nは、ルータ装置101と接続され、ルータ装置104と同様に、FANP、TDP、LDP等のプロトコルが動作しており、カットスループスの設定やカットスループスの最終段としての機能を有する。

【0041】

なお、上記のように各ルータ装置101～103/104/111～11Nについて（以下の説明のために便宜的に）区別をしたが、全てのルータ装置がカットスルー転送機能を有するようなものであっても構わない。

【0042】

図1のネットワークでは、各ルータ間において、IPデータグラムの転送先を決定するルーチングプロトコルが動作しており、動作しているプロトコルとしては、OSPF（Open Shortest Path Fast）等が考えられる。

【0043】

ネットワーク上でルーチングプロトコルを動作させることにより、ルータ装置101では、ネットワーク120やルータ装置104にデータグラムを転送する場合に、ルータ装置102を経由するルートとルータ装置103を経由するルートのいずれを採用しても同一のコストで到達可能であることを認識することができる。これによって、ルータ装置101では、データグラム転送時にルータ装置102とルータ装置103のどちらに転送しても良いことになる。

【0044】

このような構成において、ルータ装置111～11Nから、ネットワーク120またはルータ装置104に対し、FANP、TDP、LDP等のプロトコルを用いた、カットスループス131～13Nを生成する場合について説明する。

【0045】

この場合、ルータ装置101では、ルータ装置111～11Nからそれぞれ到着した設定メッセージを、ルータ装置104またはネットワーク120の方向に

向かって転送することになるが、ルータ装置 101 では、データグラムを転送する場合と同様に、カットスルーを延長するためにメッセージを転送する次のルータ装置として、ルータ装置 102 とルータ装置 103 のどちらに転送しても良いことになる。

【0046】

ここで、図 2 に、図 1 に例示したようなネットワーク構成内に存在するルータ装置 101 の一構成例を示す。なお、これと同一の構成をルータ装置 102 やルータ装置 103 が持っても良いが、図 1 のネットワーク構成例ではルータ装置 101 のみマルチパスを持つため、このルータ装置 101 を例としている。

【0047】

IP 処理部 201 は、IP データグラムの宛先情報に基づいて当該データグラムが自分宛のものである否か調べ、自分宛のデータグラムについては、受信処理を行って、上位のプロトコル（例えば、TCP）に転送し、自分宛でないものについては、次の転送先を決定し、次段のルータ装置に転送するための処理を行う。

【0048】

なお、ルータ装置 101 では、ラベルによるスイッチング機能を有するため、後者の自分宛でないデータグラムに対する転送機能を有する必要性は必ずしもない。ただし、ルータ装置 104 やルータ装置 111～11N が図 1 の構成を取る場合には、これらについては、カットスルーの始点および終点となり得るため、データグラムの転送機能は必要となる。

【0049】

カットスルー制御部 202 は、隣接するルータ装置との間でデータグラムフローおよびラベル情報について共通の認識を得る（データグラムフローとラベルの情報をやり取りする）ために使用するプロトコルを実行する部分である。

【0050】

ルーティングテーブル 203 は、宛先アドレスから次に送るべきルータ装置を求めるためのテーブルであり、次段情報は複数持つことも可能であるが、経路によってはマルチパスが存在しない場合もあるので、必ずしも複数ではない。

【0051】

ネットワークインタフェース211～21Nは、それぞれルータ装置111～11Nと接続されるもので、物理層としては、FANP、TDP、LDP等が使用可能なものであれば、どのようなものでも良い。例えば、ATM、フレームリレー、イーサネットなどが考えられる。

【0052】

ネットワークインタフェース221は、ルータ装置102に接続され、動作としてはネットワークインタフェース211～21Nと同様のものである（ここでは、ネットワークインタフェース211～21Nと同一構成とする）。

【0053】

ネットワークインタフェース222は、ルータ装置103に接続され、動作としてはネットワークインタフェース211～21Nと同様のものである（ここでは、ネットワークインタフェース211～21Nと同一構成とする）。

【0054】

スイッチ部204は、カットスルー転送が可能な場合には、ネットワークインタフェースからネットワークインタフェースへと直接スイッチング可能となるようなスイッチ装置である。

【0055】

各ネットワークインタフェースの構成について、ネットワークインタフェース211を例にとって説明する。

【0056】

物理層処理部231は、ネットワークインタフェースが収容する物理層に応じて異なる処理が行われるが、ATMの場合にはセル同期等の処理、イーサネットの場合にはMAC処理等が、物理層処理にあたる。

【0057】

ラベル処理部232は、受信したデータグラム（フレーム）のヘッダ情報よりラベルを抽出した後、ラベルテーブル233を検索することで、次のルータ装置へのラベルを決定し、IPヘッダ等を使用したデータグラム転送処理を実行せずに、スイッチ部204を通じてそのまま次のルータに転送するための処理を行う

ものである。

【0058】

なお、ATMの場合には、ラベル処理部232およびラベルテーブル233は、ATMスイッチ等で使用するVPI/VCIによるスイッチテーブルがそのまま使用可能であり、ルータ装置101が保有するネットワークインタフェースが全てATMの場合には、ネットワークインタフェース211～21N、221、222、ラベル処理部232、ラベルテーブル233、スイッチ部204の全ての機能がATMスイッチで実現可能となる。

【0059】

以下では、ルータ装置101のカットスループス設定手順について説明する。

【0060】

図3に、図1に例示したネットワークにおいてルータ装置111～11Nから最終宛先をネットワーク120またはルータ装置104とするカットスループスの生成を行ったような場合における、ルータ装置101の処理手順の一例を示す。ここでは、マージを行わない場合の例について説明する。

【0061】

なお、カットスループプロトコル生成用のプロトコルによって動作は異なるが、ここでは、カットスループ生成用のメッセージ（以下、設定メッセージと呼ぶ）を契機に次段のルータにメッセージを転送する場合の一例について説明する。

【0062】

ルータ装置101のネットワークインタフェース211～21Nより受信したカットスループ生成用のメッセージは、ラベル処理部232にてラベルテーブル223検索後、直接他のネットワークインタフェースに転送せずに、IP等のデータグラム処理を行う旨を決定し、IP処理部201に転送する（ステップS301）。

【0063】

なお、データグラム処理を行うための条件は各プロトコルによって異なるが、基本的にスイッチ部204にてIP処理部201に対し転送する処理を行う点は同様である。

【0064】

設定メッセージを転送されたIP処理部201では、自身が受信すべきメッセージがどうかの判断を行い、カットスループロトコル用のメッセージの場合には、カットスルー制御部202にデータが送られる（ステップS302）。

【0065】

一方、IP処理部201にて自身で受信すべきメッセージでないと判断した場合には、IP転送処理を行うことになり、IP処理部201は、ルーティングテーブル203を検索した後、次のルータ装置を決定し、次のルータ装置にデータを転送する（ステップS303）。なお、このとき、ルータ装置101にてIP転送機能を有さないような装置構成の場合には、データグラムは廃棄される。

【0066】

さて、IP処理部201より設定メッセージを受信したカットスルー制御部202では、受信メッセージ中に含まれるカットスルーパスを生成するための最終宛先を獲得する（ステップS304）。

【0067】

カットスルー制御部202では、獲得した最終宛先からルーティングテーブル203を検索して、次段のルータ装置を求め（ステップS305）、最終宛先が到達不能の場合には、そのメッセージの転送は停止される（ステップS306）。

【0068】

なお、メッセージの転送が停止された後の動作は、カットスルー生成プロトコルによって異なり、ルータ装置101から前段のルータ装置111～11Nに対し、応答（以下、応答メッセージを設定完了メッセージと呼ぶ）を返すものや、そのままプロトコル動作を停止しておくものが存在する。

【0069】

一方、ステップS305にて最終宛先が存在する場合には、そのルーティングテーブル203のエントリより、次段のルータ装置の情報を獲得する。

【0070】

図1のネットワーク例では既に述べたように、ルータ装置101において、ルータ装置104およびネットワーク120に対しては、ルータ装置102および

ルータ装置 103 が次段情報として保持されている。

【0071】

このとき、カットスルー制御部 202 では、同一最終宛先に対し、既に設定されているカットスルーパスの本数（カットスルー数）を確認する（ステップ S307）。

【0072】

設定済みカットスルー数は、次段情報を求めるための判断に使用し、例えば、図 1 のネットワーク構成のように、次段情報が 2 つ存在する場合には、設定済みカットスルー数が 0 および偶数のときにはルータ装置 102 を選択し、奇数のときには、ルータ装置 103 を選択するようなアルゴリズムにより、次段情報が決定される（ステップ S308）。

【0073】

なお、次段情報が n 存在する場合には、設定済みカットスルー数を c とすると、例えば、 $c \bmod n$ の値に基づいて、各次段ルータ装置を順番に選択していけばよい。

【0074】

また、この次段ルータ装置選択アルゴリズムは上記のものに限らず、各次段ルータ装置が同一（もしくはほぼ同一）のパス数に設定される、あるいはパスが各次段ルータ装置に均等に配分される、ようにする方法であれば、どのようなものでも良い。

【0075】

このようにして次段ルータを決定したカットスルー制御部 202 では、次段ルータ装置がルータ装置 102 と決定された場合にはネットワークインタフェース 221 通じてルータ装置 102 に、また、ルータ装置 103 と決定された場合にはネットワークインタフェース 222 を通じてルータ装置 103 に、該設定メッセージを送出する（ステップ S309）。

【0076】

この後、カットスルーの終点であるルータ装置 104 もしくは途中のルータより、設定完了メッセージが到着した時点で（ステップ S310）、受信したメッ

セージ中に含まれるフロー情報とラベル情報とそれに対応する送信フロー情報とラベル情報の関連付けが完了し、入力インタフェース上 211～211N 上から出力インタフェース 221 または出力インタフェース 222 へのカットスループスの設定が可能となり、カットスルー転送への移行が可能となる（ステップ S311）。

【0077】

カットスルー転送への移行は、図 2 のような構成のルータ装置の場合には、受信インタフェース 211～211N 上に存在するラベルスイッチテーブル 233 上に、出力ラベル情報を設定することで完了となる。

【0078】

カットスルー転送への移行が完了となったときに、カットスルー制御部 202 では、設定完了メッセージを前段のルータ装置に転送する。

【0079】

このように次段ルータを決定する手順において、既に設定されたカットスループスの本数を参考に同一（もしくはほぼ同一）の本数となるようにすることで、パスレベルのロードバランスを取ることが可能となる。

【0080】

次に、図 2 に示したルータ装置 101 について、スイッチ部 204 にて複数の入力ラベルを一つの出力ラベルに送出可能なマージ機能をサポートしている場合について説明する。

【0081】

図 4 に、ルータ装置 101 がマージ機能を有する（マージを行う）場合において、例えばルータ装置 111～111N が最終宛先をネットワーク 120 またはルータ装置 104 とするような、カットスループスの生成時における、ルータ装置 101 の処理手順の一例を示す。

【0082】

ルータ装置 101 のネットワークインタフェース 211～211N より受信したカットスループス生成用のメッセージは、ラベル処理部 232 にてラベルテーブル 223 検索後、直接他のネットワークインタフェースに転送せずに、IP 等の

データグラム処理を行う旨を決定し、IP処理部201に転送する（ステップS401）。

【0083】

なお、データグラム処理を行うための条件は各プロトコルによって異なるが、基本的にスイッチ部204にてIP処理部201に対し転送する処理を行う点は同様である。

【0084】

設定メッセージを転送されたIP処理部201では、自身が受信すべきメッセージがどうかの判断を行い、カットスループロトコル用のメッセージの場合には、カットスルー制御部202にデータが送られる（ステップS402）。

【0085】

一方、IP処理部201にて自身で受信すべきメッセージでないと判断した場合には、IP等のデータグラム転送処理を行うが、このときに、IP処理部201は、ルーチングテーブル203を検索した後、次のルータ装置を決定し、次のルータ装置にデータを転送する（ステップS403）。なお、このとき、ルータ装置101にてデータグラム転送機能を有さない場合には、データを廃棄しても構わない。

【0086】

さて、IP処理部201より設定メッセージを受信したカットスルー制御部202では、メッセージ中に含まれるカットスルーパスを生成するための最終宛先を獲得する（ステップS404）。

【0087】

メッセージより獲得した最終宛先からルーチングテーブル203を検索して、次段のルータ装置を求める（ステップS405）。最終宛先が到達不能の場合には、そのメッセージの転送は停止される（ステップS406）。

【0088】

なお、メッセージの転送が停止された後の動作は、カットスルー生成プロトコルによって異なり、ルータ装置101から前段のルータ装置111～11Nに対し、応答を返すものや、そのままプロトコル動作を停止しておくものが存在する

【0089】

一方、ステップS405にて最終宛先が存在する場合には、そのルーチングテーブル203のエントリより次段のルータ装置の情報を獲得可能であり、図1のルータ装置101の場合には、既に述べたようにルータ装置104およびネットワーク120に対してはルータ装置102およびルータ装置103が次段情報として保持されている。

【0090】

カットスルー制御部202では、次段ルータ装置を選択することになるが、このときに、同一最終宛先に対し、既に設定されているカットスルーパスの本数を確認する（ステップS407）。

【0091】

設定済みカットスルー数は、次段情報を求めるための判断に使用し、図3に示した場合と同様に、例えば、設定済みカットスルー数が0および偶数の場合にはルータ装置102を選択し、奇数の場合には、ルータ装置103を選択するようなアルゴリズムにより、次段情報が決定される（ステップS408）。前述と同様に、このアルゴリズムは上記のものに限らず、各次段ルータ装置が同一パス数に設定されるなどの方法であれば、どのようなものでも良い。

【0092】

次に、設定済みカットスルー数が保持される次段情報よりも少ない場合、すなわち本例では設定済みカットスルー数が0本および1本の場合には、まだカットスルーパスが最大数（本例では2本）生成されていないため、決定された次段ルータに対し、設定メッセージが送出される（ステップS409）。

【0093】

この後、カットスルーの終点であるルータ装置104もしくは途中のルータからの設定完了のメッセージが到着することで（ステップS410）、受信したメッセージ中に含まれるフロー情報とラベル情報とそれに対応する送信フロー情報とラベル情報の関連付けが完了し、入力インタフェース上211～211N上から出力インタフェース221または222へのカットスルーパスの設定が可能と

なり、カットスルー転送への移行が可能となる（ステップ S 4 1 1）。

【0094】

カットスルー転送への移行は、図 2 のような構成のルータ装置の場合には、受信インタフェース 2 1 1 ~ 2 1 N 上に存在するラベルスイッチテーブル 2 3 3 上に、出力ラベル情報を設定することで完了となる。

【0095】

また、設定済みカットスルー数が保持される次段情報の個数以上の場合、すなわち本例では設定済みカットスルー数が 2 本以上の場合には、既に送出すべきカットスルーパスはそれぞれの次段ルータとの間で設定されており、カットスルー制御部 2 0 2 では、決定された次段ルータ情報に対し、設定されたカットスルーパスにマージして送出するように、受信した設定メッセージの入力インタフェース上 2 1 1 ~ 2 1 1 N 上のラベルスイッチテーブルに設定することで完了する（ステップ S 4 1 1）。

【0096】

この後、設定メッセージを送出してきたルータ装置に対し、設定完了メッセージを応答する。

【0097】

なお、このようにマージを行う場合には、ステップ S 4 0 8 の次段情報の決定において、全ての次段情報が 1 回ずつ選択されるまでは、1 つの次段情報が複数回選択されないようにするのが好ましい（上記で示した例では、そのようになっている）。ただし、選択基準を緩やかにして、全ての次段情報が 1 回ずつ選択されるまでに、1 つの次段情報が複数回選択されても構わないようにすることも可能である。この場合には、上記で示した手順で、設定済みカットスルーパス数 \geq 保持される次段情報数の条件は、マージ対象となるカットスルーパスが既に設定されている、という条件になり、設定済みカットスルーパス数 $<$ 保持される次段情報数の条件は、マージ対象となるカットスルーパスが未だ設定されていない、という条件になる（なお、図 4 の手順においては、この条件も使用できる）。

【0098】

次に、選択可能な複数の次段情報に対応する各ネットワークインタフェースの

リンク速度（ネットワークの帯域）が異なることを考慮する場合について説明する。

【0099】

例えば、ルータ装置101のネットワークインタフェース221とネットワークインタフェース222のリンク速度が異なる場合が考えられる。従って、リンク速度を考慮して負荷分散すると好ましい。

【0100】

図5に、このような場合のカットスループスの生成時のルータ装置101の処理手順の一例を示す。ここでは、マージしない場合について説明する。

【0101】

ここでも、ルータ装置111～11Nから最終宛先をネットワーク120またはルータ装置104とするようなカットスループスの生成時を例にとるとともに、ネットワークインタフェース221とネットワークインタフェース222のリンク速度が1対2の場合（ネットワークインタフェース222が高速）を例にとる。

【0102】

ルータ装置101のネットワークインタフェース211より受信したカットスループス生成用のメッセージは、ラベル処理部232においてラベルテーブル223検索した後、自身でIP処理を行う旨を決定し、IP処理部201に転送する（ステップS501）。

【0103】

設定メッセージを転送されたIP処理部201では、自身が受信すべきメッセージがどうかの判断を行い、カットスループプロトコル用のメッセージの場合には、カットスルー制御部202にデータが送られる（ステップS502）。

【0104】

一方、IP処理部201にて自身で受信すべきメッセージでないと判断した場合には、IP等のデータグラム転送処理を行うが、このときに、IP処理部201は、ルーティングテーブル203を検索した後、次のルータ装置を決定し、次のルータ装置にデータを転送する（ステップS503）。なお、このとき、ルータ

装置 101 にて IP 転送機能を有さない場合には、データを廃棄しても構わない。

【0105】

さて、IP 処理部 201 より設定メッセージを受信したカットスルー制御部 202 では、メッセージに含まれる情報より、カットスループスを生成するための最終宛先を獲得する（ステップ S504）。

【0106】

メッセージより獲得した最終宛先からルーティングテーブル 203 を検索して、次段のルータ装置を求める（ステップ S505）。最終宛先が到達不能の場合には、そのメッセージの転送は停止される（ステップ S506）。

【0107】

なお、メッセージの転送が停止された後の動作は、カットスルー生成プロトコルによって異なり、ルータ装置 101 から前段のルータ装置 111～11N に対し、応答を返すものや、そのままプロトコル動作を停止しておくものが存在する。

【0108】

一方、ステップ S505 にて最終宛先が存在する場合には、そのルーティングテーブル 203 のエントリより次段のルータ装置の情報を獲得可能であり、図 1 のルータ装置 101 の場合には、既に述べたようにルータ装置 104 およびネットワーク 120 に対してはルータ装置 102 およびルータ装置 103 が次段情報として保持されている。

【0109】

カットスルー制御部 202 では、次段ルータ装置を選択することになるが、このときに、同一最終宛先に対し、既に設定されているカットスループスの本数を確認する（ステップ S507）。

【0110】

設定済みカットスルー数は、次段情報を求めるための判断に使用し、例えば、設定済みカットスルー数が 3 の倍数の場合にはルータ装置 102 を選択し、それ以外の場合には、ルータ装置 103 を選択するようなアルゴリズムにより、次段

情報が決定される（ステップS508）。

【0111】

なお、次段情報が n 存在し、各次段 i に対するリンク速度が $r(i)$ である場合には、設定済みカットスルー数を c とすると、例えば、 $y = c \bmod (r(1) + \dots + r(n))$ の値に基づいて、各次段ルータ装置をリンク速度に応じた頻度で選択していけばよい（次段 i を選択すべき y の値の個数（種類数）を、次段 i のリンク速度 $r(i)$ に比例させる）。

【0112】

例えば、次段情報が3つ存在し（次段1、次段2、次段3とする）、次段1に対するリンク速度：次段2に対するリンク速度：次段3に対するリンク速度＝1：2：3である場合（リンク速度の比の要素（この場合、1、2、3）の総和＝6）、 $y = c \bmod$ リンク速度の比の要素の総和（＝6）の値が0ならば次段3を選択し、1、2、3、4、5ならば、それぞれ、次段2、次段1、次段3、次段2、次段3を選択する、というような方法を用いてもよい。

【0113】

また、選択のアルゴリズムについては上記のものに制限されるものではなく、各次段ルータ装置へのパス数の比率がリンク速度の比率と一致（もしくはほぼ一致）させることができるような方法であれば、どのようなものでも良い。

【0114】

このようにして次段ルータを決定したカットスルー制御部202では、次段ルータに対し設定メッセージが、次段ルータがルータ装置102の場合にはネットワークインタフェース221を通じ、ルータ装置103の場合にはネットワークインタフェース222を通じて、送出される（ステップS509）。

【0115】

この後、カットスルーの終点であるルータ装置104もしくは途中のルータからの設定完了のメッセージが到着することで（ステップS510）、受信したメッセージ中に含まれるフロー情報とラベル情報とそれに対応する送信フロー情報とラベル情報の関連付けが完了し、入力インタフェース上211～211N上から出力インタフェース221または222へのカットスルーパスの設定が可能と

なり、カットスルー転送への移行が可能となる（ステップ S 5 1 1）。

【0 1 1 6】

カットスルー転送への移行は、図 2 のような構成のルータ装置の場合には、受信インタフェース 2 1 1 ～ 2 1 N 上に存在するラベルスイッチテーブル上に、出力ラベル情報を設定することで完了となる。

【0 1 1 7】

このように、基本的にリンク速度が異なる場合には、次段ルータ装置の選択のアルゴリズムを変更するだけで対応可能であることがわかる。

【0 1 1 8】

なお、リンク速度の相違を考慮する場合で、かつ、マージされるスイッチを使用する（マージを行う）場合についても、図 4 の例について行った説明とほとんど同様であり、ただステップ S 4 0 8 の選択処理を例えば図 5 のステップ S 5 0 8 のようにリンク速度を考慮したものに変更すればよい。

【0 1 1 9】

また、本実施形態では、図 3 ～ 図 5 の処理にて次段情報の決定時、すなわち、ステップ S 3 0 8、ステップ S 4 0 8、ステップ S 5 0 8 の処理アルゴリズムを変更することで、マージする場合についても、マージしない場合についても、様々な用途およびロードバランスを実施することが可能である。そのいくつかの応用例を以下に示す。

【0 1 2 0】

ルータ装置 1 0 1 では、ルータ装置 1 0 4 またはネットワーク 1 2 0 以外にも、様々なカットスルーパスが生成されることがある。例えば、ルータ装置 1 0 1 では、別の宛先へのカットスルーパスについて、ルータ装置 1 0 2 を次段ルータ装置として既に設定されたカットスルーパスの数とルータ装置 1 0 3 を次段ルータ装置として既に設定されたカットスルーパスの数を考慮し、どちらか一方（例えば、ルータ装置 1 0 2 とする）が多く設定されている場合には、他方のルータ装置（この場合、ルータ装置 1 0 3）への選択確率が上がるように、重み付けする。これにより、ルータ装置 1 0 2 およびルータ装置 1 0 3 に対するトータルのカットスルーパスのバランスを取ることが可能である。

【0121】

また、カットスループスの生成時にエッジルータ装置にて、QoS情報が含まれているような場合には、既にQoSを含んで設定されているカットスループス数によって、指定QoSを満足していることが容易なルータ装置を選択するように、アルゴリズムを変更すればよい。

【0122】

その他にも、種々のバリエーションが考えられる。

【0123】

次に、図6に、本実施形態におけるIP等のヘッダ情報から次段情報を求めるために使用するルーチングテーブル203の構成例を示す。

【0124】

宛先アドレス601の欄は、転送時および設定メッセージの宛先アドレス情報から次段情報を求めるためのキーとなる情報である。

【0125】

次段情報602の欄は、該当する宛先情報について保持する次段情報の数を保持するために使用し、マルチパスの場合には2以上の値が設定されている。

【0126】

次段情報611～61Nは、該当する宛先情報ヘッダグラムを転送するための次段情報のアドレスを保持するもので、次段情報602で設定された値によって、エントリ数が増える。ただし、このエントリには、実際の次段ルータ装置のアドレス情報が入っている必要はなく、次段ルータ装置情報が格納されている情報の位置を示すような情報でも構わない。

【0127】

カットスルー設定数603の欄は、該当する宛先情報について、既に設定されているカットスループスの設定数を表し、図3～図5等で示した処理について、複数存在する次段ルータ装置を決定する処理で使用する場合がある。

【0128】

図7に、本実施形態における受信したデータグラム（フレーム）に付加されるラベル値から送出インタフェースや送出ラベルを求め、IP処理を介さずに転送

するために使用するラベルスイッチテーブル 233 の構成例を示す。

【0129】

受信ラベル 701 の欄は、受信フレームのヘッダ情報に含まれるラベルと一致するもので、本テーブルの検索のキーとなるものである。ATM の場合には VPI/VCI が使用される。

【0130】

送信インタフェース 702 は、送出すべきネットワークインタフェースを指定するもので、スイッチ部 204 にてスイッチするために使用する。

【0131】

送出ラベル値欄 703 は、次のルータ装置に送るために付加するラベル、および物理レイヤが異なる場合にはその物理レイヤに関するヘッダ情報を格納するために使用される。

【0132】

以上のように本実施形態によれば、ロードバランスを実行する場合に、エッジルータの関与なしに、実際にマルチパス情報を持つルータ装置のみにて行うことができる。

【0133】

また、制御や実装の複雑化を回避し、効果的なロードバランスを実現することができる。

【0134】

さて、これまでは、カットスループスの設定、特に負荷分散のための次段の選択について説明してきたが、以下では、実際に流れるトラヒック量に応じて行うカットスループスの再設定について説明する。

【0135】

図 8 に、実際に流れるトラヒック量に応じてカットスループスの再設定を行う機能を持たせたルータ装置 101 の構成例を示す。

【0136】

IP 処理部 801 は、IP データグラム宛先情報に基づいて当該データグラムが自分宛のものである否か調べ、自分宛のデータグラムについては、受信処理

を行って、上位のプロトコル（例えば、TCP）に転送し、自分宛でないものについては、次の転送先を決定し、次段のルータ装置に転送するための処理を行う。

【0137】

なお、ルータ装置 101 では、ラベルによるスイッチング機能を有するため、後者の自分宛でないデータグラムに対する転送機能を有する必要性は必ずしもない。

【0138】

カットスルー制御部 802 は、隣接するルータ装置との間でデータグラムフローおよびラベル情報について共通の認識を得る（データグラムフローとラベルの情報をやり取りする）ために使用するプロトコルを実行する部分である。

【0139】

ルーチングテーブル 803 は、宛先アドレスから次に送るべきルータ装置を求めるためのテーブルであり、次段情報は複数持つことも可能であるが、経路によってはマルチパスが存在しない場合もあるので、必ずしも複数ではない。

【0140】

ネットワークインタフェース 811～81N は、それぞれルータ装置 111～11N と接続されるもので、物理層としては、FANP、TDP、LDP 等が使用可能なものであれば何でも良く、例えば ATM、フレームリレー、イーサネットなどが考えられる。

【0141】

ネットワークインタフェース 821 は、ルータ装置 102 に接続され、動作としてはネットワークインタフェース 811～81N と同様のものである（ここでは、ネットワークインタフェース 811～81N と同一構成とする）。

【0142】

ネットワークインタフェース 822 は、ルータ装置 103 に接続され、動作としてはネットワークインタフェース 811～81N と同様のものである（ここでは、ネットワークインタフェース 811～81N と同一構成とする）。

【0143】

スイッチ部 804 は、カットスルー転送が可能な場合には、ネットワークインタフェースからネットワークインタフェースへと直接スイッチング可能となるようなスイッチ装置である。

【0144】

各ネットワークインタフェースの構成について、ネットワークインタフェース 821 を例に取って説明する。

【0145】

物理層処理部 831 は、ネットワークインタフェースが収容する物理層に応じて異なる処理が行われるが、ATM の場合にはセル同期等の処理、イーサネットの場合には MAC 処理等が、物理層処理にあたる。

【0146】

物理層カウンタ 832 は、ネットワークインタフェース 821 から送出される物理層フレーム数を保持するものである。

【0147】

ラベル処理部 832 は、受信したデータグラム（フレーム）のヘッダ情報よりラベルを抽出した後、ラベルテーブル 833 を検索することで、次のルータ装置へのラベルを決定し、IP 処理を実行せずに、ラベルスイッチ部 804 を通じてそのまま次のルータに転送するための処理を行うものである。

【0148】

なお、ATM の場合には、ラベル処理部 833 およびラベルテーブル 834 は、ATM スイッチ等で使用する VPI/VCI によるスイッチテーブルがそのまま使用可能であり、ルータ装置 101 が全て ATM の場合には、ネットワークインタフェース 811～81N、821、822、ラベル処理部 832、ラベルテーブル 833、スイッチ部 804 の全ての機能が ATM スイッチで実現可能となる。

【0149】

実際に流れるトラフィック量に応じたカットスルーパスの再設定手順としては、概略的には、対象となるネットワークインタフェース毎にトラフィック量の計数値を定期的に参照し、所定の基準に従ってカットスルーパスを再設定すべきか否

かを判断し、再設定すべきと判断された場合に、所定の方法で選択した1または複数のカットスループスの再設定（マージが行われていない場合には設定／解放、マージが行われている場合にはラベルスイッチテーブルの内容変更）を逐次的にもしくは纏めて行う。

【0150】

図9に、図1に示したネットワークにおいて、例えばルータ装置111～11Nが最終宛先をネットワーク120またはルータ装置104とするカットスループを、図3に例示した方法等で生成した後に、実際の送出トラフィックを測定し、カットスループスの再設定を行う処理手順の一例を示す。ここでは、ルータ装置101でマージしない場合の例について示す。

【0151】

この手順は、適当なタイミングで、繰り返し実行するものとする。

【0152】

ルータ装置101では、ルーティングテーブル803でマルチパスが存在する場合（本例では、ネットワーク120およびルータ装置104）に、その出カインタフェースとして、指定されたものについて、トラフィックを測定する（ステップS901）。この場合、測定されるのはネットワークインタフェース821とネットワークインタフェース822である。

【0153】

トラフィックの測定は設定したカットスループス上のトラフィックでも可能であるが、バランスが取れているかどうかの確認のためには、ネットワークインタフェース821上に存在する物理層カウンタ832を用いた方が効率は良い。ネットワークインタフェース822についても同様である。

【0154】

ネットワークインタフェース821とネットワークインタフェース822についてそれぞれの物理層カウンタ832により送出カウンタ値を取得したならば、両者の送出カウンタ値を比較あるいは評価して、それら値の関係（例えば、比率）が許容範囲であるかどうかをチェックする（ステップS902）。

【0155】

ここで、許容範囲とは、ネットワークインタフェース 821 とネットワークインタフェース 822 のリンク速度や使用率によって決定されるもので、例えばネットワークインタフェース 821 のトラフィック量の計数値とネットワークインタフェース 822 のトラフィック量の計数値との比がそれらのリンク速度の比率に一致もしくは近似していることが好ましい。また、双方ともトラフィックが少ない場合にはバランスがとれていなくても許容範囲としても良い。

【0156】

上記のチェック後、許容範囲内の場合には、トラフィック的にはバランスが取れているものとして終了して良い（ステップ S903）。

【0157】

一方、ステップ S902 において許容範囲を超えている場合には、送出量の多いネットワークインタフェース（例えばネットワークインタフェース 821 側とする）において、マルチパスで設定されたカットスループスの中から再設定対象とするカットスループスを 1 つ選択し（ステップ S904）、該選択したカットスループスを流れるトラフィック量の計数値を調べる（ステップ S905）。

【0158】

次に、チェックしたカットスループスのトラフィック量 x を送出量の少ないネットワークインタフェース（822）のトラフィック量 M に加算した値（ $M+x$ ）と、該チェックしたカットスループスのトラフィック量 x を送出量の多いネットワークインタフェース（821）のトラフィック量 N から減算した値（ $N-x$ ）とを比較もしくは評価して（ステップ S906）、許容範囲、送出量の多いネットワークインタフェース（821）側に偏り過ぎ、送出量の少ないネットワークインタフェース（822）側に偏り過ぎをチェックする。

【0159】

許容範囲または送出量の多いネットワークインタフェース（821）側に偏っている場合には、トラフィックを測定したカットスループスの次段ルータを、送出量の多いネットワークインタフェース（821）に対応するルータ装置（102）から送出量の少ないネットワークインタフェース（822）に対応するルータ装置（103）に切り替えるようにカットスループスの再設定を行い（ステップ

S907)、そして、許容範囲の場合は終了となり(ステップS908)、送出量の多いネットワークインタフェース(821)に偏っている場合には、ステップS904以降の処理を繰り返す(ステップS909)。

【0160】

再設定方法は、カットスループロトコルによって異なり、カットスループスを解放後に設定するものと、設定後に解放処理を行うものがある。

【0161】

一方、送出量の少ないネットワークインタフェース(822)側に偏っている場合には、測定したカットスループスの出力先を変更せずに、別のカットスループスを切り替えるために、ステップS904以降の処理を繰り返す(ステップS910)。

【0162】

このような方法の場合には、バランスが許容範囲に達しない場合も考えられるが、その場合には、現状の状態と再設定後に想定される状態とで、許容範囲に近い方法を取ることも考えられる。

【0163】

次に、図10に、図1に示したネットワークにおいて、例えばルータ装置111～11Nが最終宛先をネットワーク120またはルータ装置104とするカットスループを、図4に例示した方法等で生成した後に、実際の送出トラヒックを測定し、カットスループスの再設定を行う処理手順の一例を示す。ここでは、ルータ装置101でマージする場合の例について示す。

【0164】

この手順は、適当なタイミングで、繰り返し実行するものとする。

【0165】

ルータ装置101では、ルーチングテーブル803でマルチパスが存在する場合(本例では、ネットワーク120およびルータ装置104)に、その出力インタフェースとして、指定されたものについて、トラヒックを測定する(ステップS1001)。この場合、測定されるのはネットワークインタフェース821とネットワークインタフェース822である。

【0166】

トラヒックの測定は設定したカットスループス上のトラヒックでも可能であるが、バランスが取れているかどうかの確認のためには、ネットワークインタフェース821上に存在する物理層カウンタ832を用いた方が効率は良い。ネットワークインタフェース822についても同様である。

【0167】

ネットワークインタフェース821とネットワークインタフェース822についてそれぞれの物理層カウンタ832により送出カウンタを取得したならば、両者の送出カウンタを比較あるいは評価して、それら値の関係（例えば、比率）が許容範囲であるかどうかをチェックする（ステップS1002）。

【0168】

ここで、許容範囲とは、ネットワークインタフェース821とネットワークインタフェース822のリンク速度や使用率によって決定されるもので、例えばネットワークインタフェース821のトラフィック量の計数値とネットワークインタフェース822のトラフィック量の計数値との比がそれらのリンク速度の比率に一致もしくは近似していることが好ましい。また、双方ともトラヒックが少ない場合にはバランスがとれていなくても許容範囲としても良い。

【0169】

上記のチェック後、許容範囲内の場合には、トラヒック的にはバランスが取れているものとして終了して良い（ステップS1003）。

【0170】

一方、ステップS1002において許容範囲を超えている場合には、送出量の多いネットワークインタフェース（例えばネットワークインタフェース821側とする）において、マルチパスで設定されたカットスループスの中から再設定対象とするカットスループスを1つ選択し（ステップS1004）、該選択したカットスループスを流れるトラヒック量の計数値を調べる（ステップS1005）。

【0171】

ここで、注意すべき点は、各カットスループスの送出トラヒックは既にマージ

された後であるため、マージする前の受信トラヒックを測定する必要がある。

【0172】

次に、チェックしたカットスループスのトラヒック量 x を送出量の少ないネットワークインタフェース(822)のトラヒック量 M に加算した値($M+x$)と、該チェックしたカットスループスのトラヒック量 x を送出量の多いネットワークインタフェースフェース(821)のトラヒック量 N から減算した値($N-x$)とを比較もしくは評価して(ステップS1006)、許容範囲、ネットワークインタフェース821側に偏り過ぎ、ネットワークインタフェース822側に偏り過ぎをチェックする。

【0173】

許容範囲または送出力の多いネットワークインタフェースフェース(821)側に偏っている場合には、トラヒックを測定したカットスループスのラベルスイッチテーブル823を、送出力の少ないネットワークインタフェース(822)に出力するラベルに書き換える(ステップS1007)。そして、設定後、ステップS1006の処理で許容範囲になっている場合は終了となり(ステップS1008)、送出力の多いネットワークインタフェースフェース(821)に偏っている場合には、ステップS1004以降の処理を繰り返す(ステップS1009)。

【0174】

ここで、出力先を変更したカットスループスについて、プロトコル処理が必要か否かは、使用しているプロトコルに依存する。

【0175】

一方、送出力の少ないネットワークインタフェース(822)側に偏っている場合には、測定したカットスループスの出力先を変更せずに、別のカットスループスを切り替えるためにステップS1004以降の処理を繰り返す(ステップS1010)。

【0176】

このような方法の場合には、バランスが許容範囲に達しない場合も考えられるが、その場合には、現状の状態と再設定後に想定される状態とで、許容範囲に近

い方法を取ることも考えられる。

【0177】

なお、図9の手順のステップS906や図10の手順のステップS1006における比較あるいは評価の基準としては、種々のものが考えられる。

【0178】

例えば、選択したカットスループスを再設定すると、トラフィック量の大小関係が逆転するとき、すなわち $N > M$ かつ $(N - x) < (M + x)$ となるときには、選択したカットスループスを採用しない（再設定しない）とする場合には、 $(N - x) < (M + x)$ のとき送出量の少ないネットワークインタフェース（本例では822）側に偏り過ぎとしてステップS906／ステップS1006からステップS904／ステップS1004に戻り、それ以外で許容範囲のときステップS907／ステップS1007から終了し、それ以外のとき依然として送出量の多いネットワークインタフェース（本例では821）側に偏り過ぎとしてステップS906／ステップS1006からステップS904／ステップS1004に戻るようにしてもよい。

【0179】

また、例えば、上限許容値 $\geq (N - x) / (M + x) \geq$ 下限許容値のとき許容範囲、 $(N - x) / (M + x) >$ 上限許容値のとき送出量の多いネットワークインタフェース（本例では821）側に偏り過ぎ、下限許容値 $> (N - x) / (M + x)$ のとき送出量の少ないネットワークインタフェース（本例では822）側に偏り過ぎとしてもよい。

【0180】

また、図9の手順のステップS904や図10の手順のステップS1004でのカットスループスの選択方法としても、種々のものが考えられる。

【0181】

例えば、ラベルスイッチテーブル834のエントリを1つずつ整列順に選択していてもよい。また、例えば、ラベルスイッチテーブル834のエントリをそのトラフィック量の計数値の多いものから順に選択していてもよい。

【0182】

また、図9や図10の手順例では、再設定対象とする1つのカットスループスの選択／選択したカットスループスを再設定するか否かの判断／再設定すると判断された場合の再設定／繰り返し処理を続行するか否かの判断といった一連の処理を繰り返し行うものであったが、最初に再設定対象とするカットスループスの選択を行い、後でまとめて再設定を行うようにしてもよい。

【0183】

また、各ネットワークインタフェースのトラフィック量と、マルチパスで設定された各カットスループスのトラフィック量とを考慮して、再設定すべき1または複数のカットスループスの最適な解もしくは次善の解を求めるようにしてもよい。この場合に、解が複数ある場合には、再設定すべきカットスループスの数が最小ものを選択するようにしてもよい。また、解が複数ある場合に、得られるバランスの程度や再設定すべきカットスループスの数などを総合的に勘案して解を選択するようにしてもよい。

【0184】

次に、図11に、本実施形態における受信したデータグラム（フレーム）に付加されるラベル値から送出インタフェースや送出ラベルを求め、IP処理を介さずに転送するために使用するラベルスイッチテーブル834の構成例を示す。

【0185】

受信ラベル1101の欄は、受信フレームのヘッダ情報に含まれるラベルと一致するもので、本テーブルの検索のキーとなるものである。ATMの場合にはVPI／VCIが使用される。

【0186】

送信インタフェース1102は、送出すべきネットワークインタフェースを指定するもので、スイッチ部804にてスイッチするために使用する。

【0187】

送出ラベル値欄1103は、次のルータ装置に送るために付加するラベルおよび物理レイヤが異なる場合には、その物理レイヤに関するヘッダ情報を格納するために使用される。

【0188】

受信フレーム数欄 1104 は、このラベルテーブル情報によって転送されたフレーム数を意味し、図 9 や図 10 で示したカットスループス単位のトラフィック測定時に参照される。

【0189】

本実施形態によれば、実際に流れるトラフィック量に応じてカットスループスの再設定を行うので、より効果的なロードバランスを実現することができる。

【0190】

図 1 ～図 7 で説明したルータ装置や設定手順に係る発明と、図 1, 8 ～11 で説明したルータ装置や再設定手順に係る発明は、組み合わせて実施することも可能であるし、独立して実施することも可能である。

【0191】

なお、以上の各機能は、ソフトウェアとしても実現可能である。

【0192】

また、各発明の各実施形態は、コンピュータに所定の手段を実行させるための（あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても実施することもできる。

【0193】

本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0194】

さて、以下では、図 12 ～図 20 を参照しながら、他の発明の実施の形態等について説明する。

【0195】

本発明は、例えば、一つ以上の VPN を收容するルータ装置（ラベルスイッチルータ装置）、その通信方法などに関するものである。

【0196】

（従来の技術）

ラベルスイッチの機能を有するルータから成る非プライベートネットワーク上

で動作するVPN (Virtual Private Network; 仮想的プライベートネットワーク) を構築することに関心が集まっている。そこでは、プライベートネットワークの各拠点が互いに離れている場合において、各拠点のCPE (Customer Premise Equipment) を非プライベートネットワーク上のLSP (Label Switched Path; ラベルスイッチパス) で相互接続する。設定された各LSP上には、VPNに属するパケットのみが流れ、そうでないパケットは流れない。

【0197】

非プライベートネットワーク上の端に在って、上記LSPを終端し、かつ、特定のプライベートネットワーク拠点を、そのCPEと接続することによって収容する装置をエッジルータと呼ぶ。複数の互いに異なるVPNが非プライベートネットワーク上に構築される場合もある。その中で一つのエッジルータが互いに異なるアドレス空間に属する複数のプライベートネットワークを収容する場合もある。

【0198】

エッジルータは、非プライベートネットワーク側のインタフェースを備え、かつ、収容する各々のCPEと接続する各々のプライベートネットワーク側のインタフェースを備える。

【0199】

この種のエッジルータの従来型としては、二種類のネットワーク層処理モジュール、すなわち、非プライベート側のモジュールとプライベート側のモジュールから構成されるものが知られている。

【0200】

非プライベート側のモジュールとプライベート側のモジュールの両者は互いにネットワークインタフェースで相互接続される。両者はLDP (Label Distribution Protocol) を実行することで両者間のLSPを形成する。一方、非プライベート側のLSPは、上記両者間のLSPとは互いに独立しているが、これらをスイッチングにて直接接続することで、最終的にLSPの終端インタフェースをプライベート側のモジュールに延長するようにして

いる。

【0201】

このようなLSP終端インタフェースは、バーチャルネットワークインタフェースと呼ばれる。なお、一つのエッジルータ装置が複数のVPNを収容する場合は、それ相応数のプライベート側モジュールを備える。

【0202】

このようなエッジルータ装置は、非プライベート側のモジュールとプライベート側モジュールとが分離された個別の装置としてもよいが、その場合、非プライベート側の装置がエッジルータとなり、プライベート側の装置がCPEとなる。

【0203】

(発明が解決しようとする課題)

このような従来型装置においては、LSPをプライベート側モジュールへと伸ばす目的のために、非プライベート側のモジュールとプライベート側モジュールとが互いにLDPで通信しなければならないため、以下のような問題がある。

【0204】

ひとつは、前者もしくは後者のどちらかが、一つのモジュール上に、非プライベートアドレス空間のインタフェースとプライベートアドレス空間のインタフェースとの両方を備えなければならないことである。これは、プライベート空間を一つだけ収容する場合でさえ、モジュール上のネットワークインタフェース毎にパケットフィルタリング機能を付加しなければならず、そのためにパケット転送能力が低下するという欠点がある。

【0205】

さらに、従来型装置がプライベート空間を複数収容する場合、現在一般的にネットワーク層処理の実装に使用されるソフトウェアもしくはハードウェアの技術では、パケットが属するネットワークのアドレス空間を複数識別することができないので、例えばネットワークインタフェースや経路制御テーブル等、あらゆるネットワーク構成要素にアドレス空間の識別子を付加しなければならない。すなわち、従来型装置構成では、現在の一般的なソフトウェアもしくはハードウェアに大幅な改造が必要とされ、実用化が難しいといえる。

【0206】

さらなる問題は、そのような複数アドレス空間の処理機能を前提とし、プライベート側モジュールがLDPを実行しなければならないことである。これは、プライベート側モジュールの実装が複雑かつ高価になるばかりでなく、LDPの仕様が更新された場合にはその都度プライベート部の実装も更新しなければならないため、将来への適応性に欠けるといえる。

【0207】

本発明は、上記のような従来型装置の欠点を除去し、プライベート側モジュールが簡易に実装でき、かつ、ラベルスイッチ標準の使用変更の際には非プライベート側のモジュールだけを変更すれば済むような、（エッジ）ルータ装置、その通信方法を提供することを目的とするものである。

【0208】

（発明が解決しようとする課題）

本発明（請求項15）は、プライベートアドレス空間から成るプライベートネットワークに接続されパケットを転送する手段を有するプライベート処理モジュールと、非プライベートアドレス空間から成る非プライベートネットワークに接続されパケットをラベルスイッチする手段を有する非プライベート処理モジュールと、を備えるシステム（エッジルータ装置、もしくは、エッジルータとCPEから成るシステム）における通信方法であって、前記非プライベート処理モジュールは、前記プライベート処理モジュールが接続されたプライベートネットワークと前記プライベートアドレス空間を共有する他のプライベートネットワークを収容する他のシステムが前記非プライベートネットワークに存在することを認識した場合に、当該他のシステムへ向けて前記非プライベートネットワークを経由するラベルスイッチパスを設定するためのプロトコル通信を行い、前記非プライベート処理モジュールは、設定された前記ラベルスイッチパスを前記プライベート処理モジュールまで延長し、前記プライベート処理モジュールに前記他のシステムのプライベートアドレスを通知し（前記プライベート処理モジュール内に前記他のシステムに対応するバーチャルインタフェースを生成し）、前記プライベート処理モジュールは、前記他のシステムのプライベートアドレス宛のパケット

を、延長された前記ラベルスイッチパスへ転送することを特徴とする。

【0209】

これにより、プライベート処理モジュールがラベルスイッチパスを設定するためのプロトコル通信を行わなくとも、また、非プライベート処理モジュールが各VPN (Virtual Private Network) のプライベートアドレス空間を管理しなくとも、プライベート処理モジュールを端点とするラベルスイッチパスを設定することができ、同一VPNに属するプライベートネットワーク同士を非プライベート (公衆) ネットワークを介してラベルスイッチパスで結ぶことが簡易に実現できる。

【0210】

好ましくは、前記プライベート処理モジュールと前記非プライベート処理モジュールとの間に予め複数のパス (例えば仮想コネクション) を用意しておき、前記ラベルスイッチパスの延長は、この複数のパスから選択した一つのパスと、設定された前記ラベルスイッチングパスとを、前記非プライベート処理モジュール内で対応付けることにより行い、前記他のシステムのプライベートアドレスは、選択された前記一つのパスと対応付けて通知されるようにしてもよい。

【0211】

好ましくは、前記ラベルスイッチパスの延長は、前記ラベルスイッチパスが前記他のシステムに到達するように設定されたことが確認された後に行うようにしてもよい。

【0212】

非プライベート処理モジュールがラベルスイッチパスを延長し、相手方のシステムに対応するバーチャルインタフェースをプライベート処理モジュール内に生成するタイミングは、少なくとも自システムの隣接ノードまでラベルスイッチパスが設定された後であっても良いが、これだと非プライベートネットワークの内部でラベルスイッチパスが終端してしまっている場合に、自システムのプライベート処理モジュールから転送されたパケットが行き先不明により失われてしまう可能性がある。上記のように、このタイミングを、相手方のシステムまで到達するようにラベルスイッチパスが設定された後とすれば、このようなパケットロス

を防ぐことができる。

【0213】

好ましくは、前記他のシステムのプライベートアドレスの通知は、パケットを前記他のシステムへ運ぶための前記ラベルスイッチと、パケットを前記他のシステムから運ぶための別のラベルスイッチとの双方が設定された場合に行うようにしてもよい。

【0214】

このようにすると、二つのラベルスイッチを組にして相手方のシステムと双方向通信ができる状態になったときに、相手方のシステムに対応する一つのバーチャルインタフェースがプライベート処理モジュール内に生成されることになり、プライベート処理モジュールは本発明でのバーチャルインターフェースを通常のネットワークインタフェースと同様に扱うことが可能になる。

【0215】

好ましくは、前記システムは、前記プライベートアドレス空間とは異なるプライベートアドレス空間から成る第二のプライベートネットワークに接続されパケットを転送する手段を有する第二のプライベート処理モジュールをも備えるものであり、前記非プライベート処理モジュールは、前記第二のプライベート処理モジュールが接続された第二のプライベートネットワークと前記異なるプライベートアドレス空間を共有する他の第二のプライベートネットワークを収容する他のシステムが前記非プライベートネットワークに存在することを認識した場合にも、当該他のシステムへ向けて前記非プライベートネットワークを経由するラベルスイッチパスを設定するためのプロトコル通信を行い、前記非プライベート処理モジュールは、設定された前記ラベルスイッチパスを前記第二のプライベート処理モジュールまで延長し、前記第二のプライベート処理モジュールに前記他のシステムのプライベートアドレスを通知し、前記第二のプライベート処理モジュールは、前記他のシステムのプライベートアドレス宛のパケットを、延長された前記ラベルスイッチパスへ転送するようにしてもよい。

【0216】

このように、本発明によれば、複数のプライベートアドレス空間を一つのエッ

ジシステムで収容する場合に、新たに収容するプライベートアドレス空間に対応して、プライベートアドレス空間が一つである場合と同様な構成のプライベート処理モジュールをもう一つ新たに非プライベート処理モジュールにつなげるだけで良いため、拡張性の高いエッジシステムを実現できる。

【0217】

ここで、もう少し本発明の作用効果に関連して説明すると、本発明によれば、例えば、エッジルータ装置等において、その非プライベート処理モジュールとプライベート処理モジュールとの間ではLSP設定プロトコルを実行する必要が無く、各々のネットワークモジュールが単一のネットワーク空間のみを処理すればよいことになる。非プライベート処理モジュールは、非プライベートネットワーク空間に属するネットワーク層の処理を実行し、かつ、収容VPNのためのLSPを設定するプロトコルを実行する。一方、プライベート処理モジュールは、プライベートアドレス空間に属するネットワーク層の処理を実行する。

【0218】

非プライベートネットワーク上で形成されたLSPは、一旦、非プライベート処理モジュールで終端するが、例えば、両モジュール同士を内部的に接続するバスと前記LSPとを、スイッチ部の操作などによって、実質的なLSPの端点をプライベート処理モジュールに延長する。そのとき、プライベート処理モジュールは、その端点をバーチャルネットワークインタフェースとして扱うが、そのために必要な情報は、非プライベート処理モジュール（内の例えばLSP設定プロトコル制御部）から得られ、例えば、それを、任意の下位レイヤプロトコルの拡張により、非プライベート処理モジュールがプライベート処理モジュールに対して通知するようにする。

【0219】

前記下位レイヤプロトコルは、LSP設定プロトコルの仕様には依存せず、例えば、単純に、下位レイヤのバス識別情報と、LSPにて相互接続された相手方エッジルータのプライベートアドレスとのペアを与える。

【0220】

以上により、プライベート処理モジュールがLSPを設定するためのプロトコ

ル通信を行わなくとも、また、非プライベート処理モジュールが書くVPNのプライベートアドレス空間を管理しなくとも、プライベート処理モジュールを端点とするLSPを設定することができ、同一VPNに属するプライベートネットワーク同士を非プライベートネットワークを介してLSPで結ぶことが簡単に実現できる。

【0221】

なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【0222】

また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【0223】

（発明の実施の形態）

<実施形態2-1>

図12は、本発明におけるルータ装置を使用して、一つのVPNを収容するようなネットワーク構成の一例について示したものである。

【0224】

ネットワーク1201は、VPN (Virtual Private Network) を収容するためのネットワークである。

【0225】

ルータ装置1211～121Mは、各ルータ間にて、転送するデータグラムフローの情報と下位レイヤの情報について共通の認識を得ることで、パケットヘッダ情報を見ることなく、下位レイヤ情報だけで、高速なデータグラム転送を可能とするラベルスイッチパス（以後、LSPとも呼ぶ）を生成する機能を有する装置であり、ラベルスイッチの中段としての役割を果たすものである。

【0226】

エッジルータ装置1221～122Nは、プライベートネットワークを収容することが可能なルータ装置であり、LSPの起点および終点となる機能を有するものである。

【0227】

ネットワーク1231～123Nは、それぞれエッジルータ装置1221～122Nに収容されている。また、ネットワーク1201を介して、一つの仮想ネットワーク（以後、VPNと呼ぶ）1202を形成している。

【0228】

ここで、VPNにおけるアドレス空間に関して説明しておく。

【0229】

VPNを構成する各プライベートネットワークでは、そのVPN内でのみ有効なローカルなアドレス空間（以後、プライベートアドレス空間と呼ぶ）が共有して用いられる。一つのプライベートアドレス空間の中では、アドレスの重複が無い。一方、非プライベートネットワークでは、複数のVPN間で有効なアドレス空間（例えば、グローバル空間もしくはプロバイダ空間）が用いられる。

【0230】

図12の例では、ネットワーク1201がグローバルアドレス空間を用いるものとし、一方、ネットワーク1231～123Nがプライベートアドレス空間を用いたVPNを構成するものとして、以後の説明を行う。

【0231】

エッジルータ装置1221～122Nは、グローバルアドレス空間とプライベートアドレス空間との境界に存在しており、それぞれのアドレス空間につきそれぞれのネットワークインタフェースを備える。

【0232】

図12の例では、非プライベートネットワーク1201側のネットワークインタフェース（1251～125N）にはグローバルアドレス空間に属するアドレスが付き、一方、プライベートネットワーク側のネットワークインタフェース（1261～126N）にはプライベートアドレス空間に属するアドレスが付く。以後、このプライベートアドレス空間には“1202”という識別子が与えられ

ることとする。

【0233】

ラベルスイッチパス1242～124Lは、VPN1202を構成するために、エッジルータ装置1221～122Nによって設定されたものである。

【0234】

LSPを流れるデータグラムは途中のルータ装置（例えば1211～121M）ではIP処理されない。したがって、そのパケットヘッダにプライベートアドレスがついていても、LSPで接続されるエッジルータ装置には直接届けられる。これを利用して、各プライベートネットワーク同士が、非プライベートネットワーク上のLSPを介して、プライベートパケットをやりとりすることが可能になる。

【0235】

例えば、図12のVPNでは、LSP1241～124Lの内部にはプライベートアドレス空間1202に属するパケットを流すようにする。これは、当該LSPの起点であるエッジルータ装置1221～122Nによって制御される。

【0236】

図13に、図12に示したエッジルータ装置について、エッジルータ1221を例にとり、その一構成例を示す。

【0237】

図13のグローバル処理モジュール1301は、非プライベートネットワークに接続するためのものである。

【0238】

IP処理部1302は、ネットワーク層の処理を行うが、ここではグローバルアドレス空間に属するパケットを処理する。

【0239】

ルーティングテーブル1303は、グローバルアドレス空間に属する宛先アドレスや次段情報を管理する。

【0240】

LSP制御部1304は、非プライベートネットワークを経由するラベルスイ

ッチパスを設定するためのプロトコル（以後、LSP設定プロトコルと呼ぶ）を実行する。

【0241】

ネットワークインタフェース1305は、非プライベートネットワークに接続するためのインタフェースであり、グローバルアドレス空間に属するアドレスが付き、前記LSP設定プロトコルの通信インタフェースとしても機能する。

【0242】

本ネットワークインタフェース1305の下位レイヤとしては、LSPが設定可能なものであれば、どのようなものでも良い。例えばATM、フレームリレー、イーサネットなどが考えられる。スイッチ部1306がこれを提供している。

【0243】

スイッチ部1306は、データリンクにてプライベート処理モジュール1311と接続されている。

【0244】

プライベート処理モジュール1311は、プライベートネットワークを収容するためのものである。

【0245】

IP処理部1312は、IP処理部1302と同様の処理を行うが、ここではプライベートアドレス空間に属するパケットを処理する。

【0246】

ルーティングテーブル1313は、プライベートアドレス空間に属する宛先アドレスや次段情報を管理する。

【0247】

ネットワークインタフェース1314は、プライベートネットワークを収容するためのインタフェースの一例であり、プライベートアドレス空間に属するアドレスが付く。

【0248】

プライベート処理モジュール1311は、一つのネットワークインタフェース1314を備えているが、もし前記プライベート空間と同一空間に属するプライ

プライベートネットワークが複数あれば、それらを収容するために、同様なネットワークインタフェースを本モジュール内に相応数備えても構わない。

【0249】

前記ネットワークインタフェースの下位レイヤとしては、必ずしもLSPが設定可能なものでなくても良い。一般的にネットワーク層の下で利用されている任意の下位レイヤにて実施可能である。

【0250】

プライベート処理モジュール1311とグローバル処理モジュール1301とを接続するデータリンク上には、あらかじめ複数のパスを用意しておく。パスの実現例としては、該データリンクが例えばATMならば仮想コネクションにて実現できる。

【0251】

各パスの識別子は、図13の例では1321～132M、1331～133Mのようになる。

【0252】

プライベート処理モジュール1311では、図12のようなLSPを張るためにLSP設定プロトコルを行わなくとも良いため、グローバル処理モジュール1301とは異なり、LSP制御部1304を実装しなくとも良い。

【0253】

以上のように、各モジュールは、それぞれがネットワーク層の処理を行うが、一つのモジュールが複数のアドレス空間を管理しなくとも良い。

【0254】

このようなネットワーク層の処理は、例えば、現在一般的なルータに使用されているソフトウェアもしくはハードウェアを組み込んで実装できる。また、そのソフトウェアもしくはハードウェアは、ネットワーク層につき一つのアドレス空間に属する処理のみを行うものでも構わない（充分である）。

【0255】

図14では、各プライベートネットワークを非プライベートネットワーク上のLSPで相互接続してVPNを構築するためのエッジルータ装置の処理手順を、

図12に示したエッジルータ装置1221の場合を例にとって説明する。

【0256】

はじめに、エッジルータ装置1221は自分自身の初期設定を行う（ステップS1401）。

【0257】

具体的には、グローバルアドレスと、プライベートアドレスならびにプライベートアドレス空間の識別子（以後、VPN識別子と呼ぶ）とをそれぞれ設定する。VPN識別子は、この例では“1202”とする。

【0258】

かかる設定の際、図13の装置構成例においては、グローバル側のアドレスはグローバル処理モジュール1301のネットワークインタフェース1305に付加され、プライベート側のアドレスはプライベート処理モジュール1311のネットワークインタフェース1314に付加される。

【0259】

さらに、これらのアドレスとVPN識別子とがLSP制御部1304に記憶される。

【0260】

ここで設定に使った（VPN識別子、プライベートアドレス、グローバルアドレス）の三つ組を、VPNメンバーシップ情報と呼ぶことにする。

【0261】

次に、図12のエッジルータ装置1221は、LSPを設定すべき相手方のシステムの存在、すなわち、自分と同じVPN識別子“1202”を有するプライベートネットワークを収容するエッジルータ装置が自分以外にも非プライベートネットワーク1201上に存在することを認識すべく、前記VPNメンバーシップ情報を追加獲得する（ステップS1402）。

【0262】

非プライベートネットワーク1201上での他者の存在を認識しなければならないため、前記メンバーシップ情報の収集は、グローバル処理モジュール1301によって実行される。かかる具体的手段としては、その際にはメンバーシップ

情報を第三者によって偽られたりしないように認証等の付加的機能を持った手段が適切であるが、例えばSNMP (Simple Network Management Protocol) やLDAP (Lightweight Directory Access Protocol) 等の動的な配布プロトコルや、もしくは手設定などが考えられる。

【0263】

グローバル処理モジュール1301は、収集したVPNメンバーシップ情報をローカルなデータベースとして保持する。

【0264】

次に、LSP制御部1304は、当該エッジルータ装置へ向けてLSPを形成するために、グローバル側のインタフェース1305からLSP設定要求メッセージを送信する(ステップS1403)。この場合は、送信用LSPの発呼を意味する。

【0265】

前記LSP設定要求メッセージにおいて、フローの識別情報には、

(VPN識別子、LSP起点エッジルータのグローバルアドレス、LSP終点エッジルータのグローバルアドレス)

の三つ組の値を含ませるようにする。この三つ組情報を含むようなフロー識別情報を、VPNフロー識別子と呼ぶことにする。

【0266】

しかる後、前記LSP設定メッセージにて要求したLSPに関するLSP設定完了通知が、図13に示したLSP処理部1304に到着すると(ステップS1404)、グローバル処理モジュール1301は、スイッチ部1306を操作し、設定された前記LSPをプライベート処理モジュール1311まで延長接続する(ステップS1405)。

【0267】

例えば、エッジルータ1221からエッジルータ122Nへ到達するLSPが設定されたことを確認したときは、ラベル値136Nを持つLSPおよび識別子の値132Mを持つパスの両者を、スイッチ部1306で接続する。

【0268】

なお、この操作は、エッジルータ装置が内部的に行うことであり、非プライベートネットワーク上のLSP接続状態には影響を与えない。

【0269】

ここで、前記ラベル値はLSP設定プロトコルにより割り当てられたものである。

【0270】

また、前記パスの識別子は、先述のあらかじめ用意されていた複数のパスから選択した一つの識別子である。

【0271】

以上の操作を行った後、グローバル処理モジュール1301は、図13に示したバーチャルインタフェース設定命令を発行し、VPN識別子に当該のプライベート処理モジュール1311内にバーチャルインタフェースを設定する（ステップS1406）。

【0272】

この場合のバーチャルインタフェース設定命令では、プライベート処理モジュール1311が備えるルーティングテーブルに新たなエントリを追加記入し得る情報を通知する。

【0273】

ここでは、バーチャルインタフェースの初期状態としてDOWN状態としておく。

【0274】

なお、グローバル処理モジュールがLSPを延長し、相手方のエッジルータに対応するバーチャルインタフェースをプライベート処理モジュール内に設定するタイミングは、少なくとも自エッジルータの隣接ノードまでラベルスイッチパスが設定された後であっても良いが、これだと非プライベートネットワークの内部でLSPが終端してしまっている場合に、自エッジルータのプライベート処理モジュールから転送されたパケットが行き先不明により失われてしまう可能性がある。上記のように、このタイミングを、相手方のエッジルータまで到達するよう

にラベルスイッチパスが設定された後とすれば、このようなパケットロスを防ぐことができる。

【0275】

しかる後、グローバル側のネットワークインタフェース1305にてLSP設定メッセージが到着する（ステップS1407）。

【0276】

このとき、まず、エッジルータ1221はVPNメンバーシップ情報データベースを照合し、前記設定メッセージを受け付けるかどうかを判断する（ステップS1408）。

【0277】

受け付けても良い場合としては、前記メッセージにてLSPのラベルに対応付けられているVPNフロー識別子について、VPN識別子と自らが収容しているプライベートネットワークのVPN識別子（この例では“1202”）とが一致し、LSP終点エッジルータのグローバルアドレスが自分自身のグローバルアドレスと一致し（以上は受信用LSPの着呼を意味する）、かつ、該起点エッジルータのプライベートアドレス（これは、この時点で持っているVPNメンバーシップ情報データベースから検索できる）との間でまだ受信用LSPが張られていない、という条件がすべて合致するときに挙げられる。

【0278】

このような場合は設定完了通知メッセージを送出し（ステップS1409）、そうでない場合は終了する（ステップS1410）。LSP設定プロトコルの仕様にも依存するが、もしNackを返す機構が定義されていれば、単に終了とせずに、Nackを返しても良い。

【0279】

設定完了通知を送出したならば、図13に示したグローバル処理モジュール1301は、スイッチ部1306を操作し、設定された前記LSPをプライベート処理モジュール1311まで延長接続する（ステップS1411）。

【0280】

例えば、エッジルータ122Nからエッジルータ1221へ到達するVPN1

202に関するLSPの設定を受け付けたときは、ラベル値137Nを持つLSPおよび識別子の値133Mを持つパスの両者を、スイッチ部1306で接続する。

【0281】

この操作の結果、LSPの実質的な終点はプライベート処理モジュール1311内のバーチャルネットワークインタフェース1341となる。

【0282】

なお、この操作は、エッジルータ装置が内部的に行うことであり、非プライベートネットワーク上のLSP接続状態には影響を与えない。

【0283】

また、前記ラベル値はLSP設定プロトコルにより割り当てられたものであり、前記パスの識別子は、先述のあらかじめ用意されていた複数のパスから選択した一つの識別子である。

【0284】

前記LSPの延長操作を行った後、グローバル処理モジュール1301は、図13に示したバーチャルインタフェース設定命令を発行し、当該のバーチャルインタフェースをUP状態にする（ステップS1412）。

【0285】

このようにすると、二つのLSPを組にして相手方のシステムと双方向通信ができる状態になったときに、相手方のシステムに対応するバーチャルインタフェースの使用が開始されることになり、プライベート処理モジュールは本発明でのバーチャルネットワークインタフェースを通常のネットワークインタフェースと同様に扱うことが可能になる。

【0286】

なお、図14には省略したが、VPNメンバーシップ配布プロトコルあるいは手設定によって、自分または相手方のエッジルータが該VPNメンバーシップから脱退するという割り込みイベントが発生した時は、いかなる状態にあっても、ただちにLSPを解放し、バーチャルインタフェースそのものを消去し、初期状態へ戻る。

【0287】

以上のような状態遷移は、すべて、LSP設定プロトコルを司るグローバル部モジュールもしくはメンバーシップ情報の変化によって駆動される。したがって、図13の装置構成例においては、バーチャルインタフェースに関する制御は、すべてグローバル処理モジュール1301からの命令によって実現可能である。

【0288】

また、プライベート処理モジュール1311は、それ自身がLSP設定プロトコルを実行する必要がなく、バーチャルインタフェースの制御についてはグローバル処理モジュール1301からの命令を待っているだけで良いので、該プライベート空間の packets 転送や経路制御のためだけに処理能力を集中することができる。

【0289】

なお、前記バーチャルインターフェース制御命令は、任意の下位レイヤプロトコルの拡張により実施可能である。

【0290】

このプロトコルは、LSP設定プロトコルの仕様には依存せず、単純に、該LSPに継ぎ足されたパスの下位レイヤ情報と、LSPにて相互接続された相手方のエッジルータのプライベートアドレスとのペアとを情報として含んでいれば良い。

【0291】

これまでに述べた処理手順がすべて行われると、図12に示したようにエッジルータ同士がLSPで相互接続され、お互いがプライベートアドレス空間1202での通常の通信ができるようになる。

【0292】

その後は、通常のネットワークと同様の方法で経路制御プロトコルを実行することができ、各プライベートネットワーク1231～123Nの間で、プライベートアドレス空間1202に属するネットワーク通信ができるようになる。

【0293】

また、これまで述べた手順により、例えば、エッジルータ1221では、プラ

イベント処理モジュール 1311 が扱うすべてのインタフェースは、プライベートアドレス空間 1201 に属することが保証されている。

【0294】

したがって、プライベート処理モジュール 1311 は、バーチャルインタフェースも含め、通常のネットワークと同様の手法で、ネットワークインタフェース間のパケット転送を行うことができる。

【0295】

このようなエッジルータ装置は、非プライベート処理モジュールとプライベート処理モジュールとが分離された個別の装置としても良いが、その場合、非プライベート処理モジュールがエッジルータ装置となり、プライベート処理モジュールが CPE 装置となる。

【0296】

なお、ここでは、双方向ともラベルスイッチパスが設定された場合にバーチャル I/F をプライベート処理モジュール内に生成する例を示したが、どちらか一方のラベルスイッチパスしか設定されていない場合でも、そのことをプライベート処理モジュールが認識できるようにしつつバーチャル I/F を生成するように変形実施することも可能である。

【0297】

図 15 は、グローバル処理モジュールによって保持される VPN メンバーシップ情報のデータベースの一例を示す。

【0298】

例えば、図 12 におけるプライベートネットワーク 123N に関する VPN メンバーシップ情報を得たとすると、図 15 のデータベースに追加されるエントリとしては、VPN 識別子の欄 1501 の値、プライベートアドレスの欄 1502 の値、グローバルアドレスの欄 1503 の値は、それぞれ (1202、126N、125N) となる。

【0299】

図 16 は、LSP 制御部 1304 によって管理される、VPN フロー識別子に関するテーブルの一形式である。

【0300】

VPNフロー識別子の欄1601は、そのVPN識別子が自ら収容するプライベートネットワークのVPN識別子と一致するもののみを記入するものとする。

【0301】

LSPのラベル値の欄1602には、LSP設定プロトコルによって前記VPNフロー識別子に割り当てられたラベル値を記入する。

【0302】

パスの識別子の欄1603には、非プライベートネットワーク上の該LSPに接続延長されたパスの識別子を記入する。

【0303】

図17に、バーチャルインタフェースの一構成例を示す。

【0304】

図17のネットワーク層および下位レイヤの入出力1701～1704は、たとえば一般的なネットワークオペレーティングシステム実装におけるソフトウェア関数として実装しても良いし、あるいはハードウェアにて実装しても良い。

【0305】

送信方向のパス1705と受信方向のパス1706とが取りまとめられ、送受信可能な一つのネットワークインタフェース1707を構成している。

【0306】

このように、バーチャルネットワークインタフェースは、LSP技術に依存しない通常のネットワークインタフェースに用いられる要素だけで構成して良い。

【0307】

送信方向のパスが送信用のLSPと接続されることにより、LSPの実質的な起点は本バーチャルネットワークインタフェースとなる。

【0308】

受信方向のパスは、受信用のLSPと接続される。

【0309】

図18に、プライベート処理モジュールが管理するルーティングテーブルの形式の一例を示す。

【0310】

バーチャルインタフェース設定命令が発行された際には、宛先の欄 1801 には相手方エッジルータのプライベートアドレスが、次段情報の欄 1802 にはあらかじめ用意していた複数のパスから選択された一つのパスの識別子の値が、状態の欄 1803 には初期状態として DOWN 状態が、インタフェース名の欄 1804 にはバーチャルインタフェースの名前が、それぞれ記入される。

【0311】

ただし、インタフェース名 1804 の値は、必ずしも前記設定命令で具体的に通知しなくとも良い。

【0312】

例えば、プライベート処理モジュールが、あらかじめローカルに管理するバーチャルインタフェース名と前記パスの識別子とを対応付けておくことで、前記設定命令で通知されたパスの識別子から一意に導けば良い。

【0313】

<実施形態 2-2>

図 19 は、本発明におけるルータ装置を使用して、複数 (N 個) の VPN を収容するようなネットワーク構成の一例について示したものである。

【0314】

ネットワーク 1901 は、VPN (Virtual Private Network) を収容するためのネットワークである。

【0315】

ルータ装置 1911 ~ 191M は、実施形態 2-1 で説明したルータ装置 1211 ~ 121M と同様の機能を有する装置である。

【0316】

エッジルータ装置 1920 ~ 192N は、実施形態 2-1 で説明した装置 1221 ~ 122N と同様の機能を有するものである。

【0317】

なお、エッジルータ装置 1920 のように、一つのエッジルータが複数 (N 個) のプライベートネットワーク 1931 ~ 193N を収容することがでいる点で

、実施形態2-1の場合とは異なっている。

【0318】

しかも、プライベートネットワーク1931~193Nは、互いに異なるVPN1941~194Nを形成している。VPNが異なれば、そのプライベートアドレス空間も互いに異なる。以後、これらアドレス空間はそれぞれ“1941”、“1942”、…、“194N”という識別子で識別されるものとする。

【0319】

また、プライベートネットワーク1991~199Nは、それぞれエッジルータ装置1921~192Nに収容されているおり、それらのアドレス空間はそれぞれ“1941”、“1942”、…、“194N”という識別子で識別されるものとする。

【0320】

エッジルータ装置1920~192Nは、グローバルアドレス空間とプライベートアドレス空間との境界に存在しており、それぞれのアドレス空間につきそれぞれのネットワークインタフェースを備える。

【0321】

エッジルータ装置1920については、非プライベートネットワーク1901側のネットワークインタフェース1950にはグローバルアドレス空間に属するアドレスが付き、一方、プライベートネットワーク側のインタフェース1951~195Nにはそれぞれプライベートアドレス空間1941~194Nに属するアドレスが付く。

【0322】

ラベルスイッチパス1961~196Nは、それぞれVPN1941~194Nを構成するために、エッジルータ装置1920~192Nによって設定されたものである。

【0323】

図20に、図19に示したエッジルータ装置について、エッジルータ装置1920を例にとり、その一構成例を示す。

【0324】

本装置は、実施形態 2-1 で述べたものと同様の機能を有する (N+1) 個のモジュールから構成される。

【0325】

図 20 のグローバル処理モジュール 2001 は、非プライベートネットワークに接続するためのものである。

【0326】

IP 処理部 2002 は、ネットワーク層の処理を行うが、ここではグローバルアドレス空間に属するパケットを処理する。

【0327】

ルーティングテーブル 2003 は、グローバルアドレス空間に属する宛先アドレスや次段情報を管理する。

【0328】

LSP 制御部 2004 は、非プライベートネットワークを経由する LSP 設定プロトコルを実行する。

【0329】

ネットワークインタフェース 2005 は、非プライベートネットワークに接続するためのインタフェースであり、グローバルアドレス空間に属するアドレスがつき、前記 LSP 設定プロトコルの通信インタフェースとしても機能する。

【0330】

本ネットワークインタフェース 2005 の下位レイヤとしては、LSP が設定可能なものであれば、どのようなものでも良い。例えば、ATM、フレームリレー、イーサネットなどが考えられる。スイッチ部 2006 がこれを提供している。

【0331】

スイッチ部 2006 は、データリンクにてプライベート処理モジュール 2011 ~ 201N と接続されている。

【0332】

プライベート処理モジュール 2011 ~ 201N は、プライベートネットワークを収容するためのものである。

【0333】

これらプライベート処理モジュールは、実施形態2-1で述べたものと同様の機能を有するが、ここでは互いに異なるプライベートアドレス空間1941～194Nに属するネットワーク層の処理を行う。

【0334】

プライベート処理モジュールとグローバル処理モジュールとを接続するデータリンク上には、あらかじめ複数のパスを用意しておく。パスの実現例としては、前記データリンクが例えばATMならば仮想コネクションにて実現できる。

【0335】

各パスの識別子は、図20の例では、2021～202M、2031～203Mのようになる。

【0336】

プライベート処理モジュール2011～201Nでは、図19のようなLSPを張るためにLSP設定プロトコルを行わなくとも良いため、グローバル部2001とは異なり、LSP制御部を実装しなくとも良い。

【0337】

非プライベートネットワーク上のLSPを利用した複数のVPNを構築するための処理手順は、図14で示した処理手順をそのまま利用することができる。

【0338】

以下では、エッジルータ装置1920の場合を例にとって説明する。

【0339】

はじめに、エッジルータ装置1920は、自分自身の初期設定を行う（ステップS1401）。

【0340】

具体的には、グローバルアドレスと、N個のプライベートアドレスならびにプライベートアドレス空間の識別子（以後、VPN識別子と呼ぶ）とをそれぞれ設定する。VPN識別子は、この例では“1941”，“1942”，…，“194N”とする。

【0341】

かかる設定の際、図 20 の装置構成例においては、グローバル側のアドレスはグローバル処理モジュールのネットワークインタフェース 2005 に付加され、N 個のプライベート側のアドレスは各々のプライベート処理モジュールのネットワークインタフェースに付加される。

【0342】

さらに、これらのアドレスと VPN 識別子とが LSP 制御部 2004 に記憶される。

【0343】

ここで設定に使った (VPN 識別子、プライベートアドレス、グローバルアドレス) の三つ組を、VPN メンバーシップ情報と呼ぶことにする。

【0344】

次に、図 19 のエッジルータ装置 1920 は、LSP を設定すべき相手方のシステムの存在、すなわち、自分と同じ VPN 識別子 1941 ~ 194N のうちのどれかに一致するプライベートネットワークを収容するエッジルータ装置が自分以外にも非プライベートネットワーク 1901 上に存在することを認識すべく、前記 VPN メンバーシップ情報を追加獲得する (ステップ S1402)。

【0345】

非プライベートネットワーク 1901 上での他者の存在を認識しなければならないため、前記メンバーシップ情報の収集はグローバル処理モジュール 2001 によって実行される。かかる具体的手段としては、実施形態 2-1 で述べたものと同様のものが考えられる。

【0346】

グローバル処理モジュール 2001 内は、収集した VPN メンバーシップ情報を実施形態 2-1 の場合と同様なローカルなデータベースとして保持する。

【0347】

次に、LSP 制御部 2004 は、当該エッジルータ装置へ向けて LSP を形成するために、グローバル側のインタフェース 2005 から LSP 設定要求メッセージを送信する (ステップ S1403)。この場合は送信用 LSP の発呼を意味する。

【0348】

前記LSP設定要求メッセージにおいて、フローの識別情報には（VPN識別子、LSP起点エッジルータのグローバルアドレス、LSP終点エッジルータのグローバルアドレス）の三つ組の値を含ませるようにする。この三つ組情報を含むようなフロー識別情報を、VPNフロー識別子と呼ぶことにする。

【0349】

VPNフロー識別子は、LSP制御部2004により、例えば図16に示すようなテーブルのエントリ形式にて管理される。この管理の方法は、実施形態2-1と同様である。

【0350】

しかる後、前記LSP設定メッセージにて要求したLSPに関するLSP設定完了通知が、図20に示したLSP処理部2004に到着すると（ステップS1404）、グローバル処理モジュール2001は、スイッチ部2006を操作し、設定された前記LSPをプライベート処理モジュール2011まで延長接続する（ステップS1405）。

【0351】

例えば、エッジルータ装置1920からエッジルータ装置192Nへ到達するLSPが設定されたことを確認したときは、ラベル値206Nを持つLSPおよび識別子の値202Mを持つパスの両者を、スイッチ部2006で接続する。

【0352】

なお、この操作は、エッジルータ装置が内部的に行うことであり、非プライベートネットワーク上のLSP接続状態には影響を与えない。

【0353】

ここで、前記ラベル値はLSP設定プロトコルにより割り当てられたものである。

【0354】

また、前記パスの識別子は、先述のあらかじめ用意されていた複数のパスから選択した一つの識別子であり、これは図16におけるパスの識別子の欄1603に記入する。

【0355】

以上の操作を行った後、グローバル処理モジュール2001は、図20に示したバーチャルインタフェース設定命令を発行し、VPN識別子に当該のプライベート処理モジュール2011内にバーチャルインタフェースを設定する（ステップS1406）。

【0356】

バーチャルインタフェースの実装やバーチャルインタフェース設定命令については、実施形態2-1と同様にして実施できる。

【0357】

しかる後、グローバル側のネットワークインタフェース2005にてLSP設定メッセージが到着する（ステップS1407）。

【0358】

このとき、まず、エッジルータ装置1920は、VPNメンバーシップ情報データベースを照合し、前記設定メッセージを受け付けるかどうかを判断する（ステップS1408）。

【0359】

受け付けても良い場合としては、前記メッセージにてLSPのラベルに対応付けられているVPNフロー識別子について、VPN識別子と自らが収容しているプライベートネットワークのVPN識別子（この例では、“1940”，“1941”，…，“194N”のどれか）とが一致し、LSP終点エッジルータのグローバルアドレスが自分自身のグローバルアドレスと一致し（以上は受信用LSPの着呼を意味する）、かつ、該起点エッジルータのプライベートアドレス（これは、この時点で持っているVPNメンバーシップ情報データベースから検索できる）との間でまだ受信用LSPが張られていない、という条件がすべて合致するときに挙げられる。

【0360】

このような場合は設定完了通知メッセージとしてAckを送出し（ステップS1409）、そうでない場合は終了する（ステップS1410）。LSP設定プロトコルの仕様にも依存するが、もしNackを返す機構が定義されていれば、

単に終了とせずに、N a c k を返しても良い。

【0361】

設定完了通知を送出したならば、図20に示したグローバル処理モジュール2001は、スイッチ部2006を操作し、設定された前記LSPを当該プライベート処理モジュールまで延長接続する（ステップS1411）。

【0362】

例えば、エッジルータ装置192Nからエッジルータ装置1920へ到達するVPN194Nに関するLSPの設定を受け付けたときは、ラベル値207Nを持つLSPおよび識別子の値203Mを持つパスの両者を、スイッチ部2006で接続する。

【0363】

この操作の結果、LSPの実質的な終点はプライベート処理モジュール201N内のバーチャルネットワークインタフェースとなる。

【0364】

なお、この操作は、エッジルータ装置が内部的に行うことであり、非プライベートネットワーク上のLSP接続状態には影響を与えない。

【0365】

また、前記ラベル値はLSP設定プロトコルにより割り当てられたものであり、前記パスの識別子は、先述のあらかじめ用意されていた複数のパスから選択した一つの識別子である。以上の情報は、図16における一つのエントリとして記入する。

【0366】

前記LSPの延長操作を行った後、グローバル処理モジュール2001は、図20に示したバーチャルインタフェース設定命令を発行し、当該のバーチャルインタフェースをUP状態にする（ステップS1412）。

【0367】

このようにすると、二つのLSPを組にして相手方のシステムと双方向通信ができる状態になったときに、相手方のシステムに対応するバーチャルインタフェースの使用が開始されることになり、プライベート処理モジュールは本発明での

バーチャルネットワークインタフェースを通常のネットワークインタフェースと同様に扱うことが可能になる。

【0368】

なお、図14には省略したが、VPNメンバーシップ配布プロトコルあるいは手設定によって、自分または相手方のエッジルータが該VPNメンバーシップから脱退するという割り込みイベントが発生した時は、いかなる状態にあっても、ただちにLSPを解放し、バーチャルインタフェースそのものを消去し、初期状態へ戻る。

【0369】

以上のような状態遷移は、すべて、LSP設定プロトコルをつかさどるグローバル部モジュールもしくはメンバーシップ情報の変化によって駆動される。したがって、図20の装置構成例においては、バーチャルインタフェースに関する制御は、すべてグローバル処理モジュールから2001の命令によって実現可能である。

【0370】

また、プライベート処理モジュール2011～201Nは、それ自身がLSP設定プロトコルを実行する必要がなく、バーチャルインタフェースの制御についてはグローバル処理モジュールからの命令を待っているだけで良いので、該プライベート空間の packets 転送や経路制御のためだけに処理能力を集中することができる。

【0371】

なお、前記バーチャルインターフェース制御命令は、任意の下位レイヤプロトコルの拡張により実施可能である。

【0372】

このプロトコルは、LSP設定プロトコルの仕様には依存せず、単純に、該LSPに継ぎ足されたパスの下位レイヤ情報と、LSPにて相互接続された相手方のエッジルータのプライベートアドレスとのペアとを情報として含んでいれば良い。

【0373】

これまでに述べた処理手順がすべて行われると、図 19 に示したようにエッジルータ同士が LSP で相互接続され、お互いが固有のプライベートアドレス空間にて通常の通信ができるようになる。

【0374】

その後は、通常のネットワークと同様の方法で経路制御プロトコルを実行することができ、プライベートネットワーク 1991 とプライベートネットワーク 1931 との間で、あるいは同様にプライベートネットワーク 199N とプライベートネットワーク 193N との間でも、各々が固有のプライベートアドレス空間にてネットワーク通信ができるようになる。

【0375】

勿論、一つのプライベートアドレス空間を共有するプライベートネットワークが複数存在する場合でも同様である。

【0376】

また、これまで述べた手順により、例えばエッジルータ 1920 では、各プライベート処理モジュールが扱うすべてのインタフェースは、そのモジュール内で単一のプライベートアドレス空間に属することが保証されている。

【0377】

したがって各プライベート処理モジュールは、バーチャルインタフェースも含め、通常のネットワークと同様の手法で、ネットワークインタフェース間のパケット転送を行うことができる。

【0378】

このようなエッジルータ装置は、非プライベート処理モジュールとプライベート処理モジュールとが分離された個別の装置としても良いが、その場合、非プライベート処理モジュールがエッジルータ装置となり、プライベート処理モジュールが CPE 装置となる。

【0379】

なお、ここでは、双方向ともラベルスイッチパスが設定された場合にバーチャル I/F をプライベート処理モジュール内に生成する例を示したが、どちらか一方のラベルスイッチパスしか設定されていない場合でも、そのことをプライベ

ート処理モジュールが認識できるようにしつつバーチャル I/F を生成するように変形実施することも可能である。

【0380】

なお、以上の各機能は、ソフトウェアとしても実現可能である。

【0381】

また、各発明の各実施形態は、コンピュータに所定の手段を実行させるための（あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても実施することもできる。

【0382】

本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0383】

（発明の効果）

本発明によれば、各プライベート部モジュールがラベルスイッチ制御やパケットフィルタリングを行わなくとも良いので、各モジュールの実装が簡易になり、パケット転送性能を損なうことなくプライベートな通信路を実現できるので、装置あたりの性能費用比の点で有効である。

また、ラベルスイッチング標準の仕様変更の際にはグローバル部モジュールのみを変更すれば済むことと、複数のプライベートアドレス空間を一つのエッジシステムで収容する場合に、新たに収容するプライベートアドレス空間に対応して、プライベートアドレス空間が一つである場合と同様な構成のプライベート処理モジュールをもう一つ新たに非プライベート処理モジュールにつなげるだけで良いことから、拡張性の高いエッジシステムを実現できる。

【0384】

さて、以下では、図 21～図 24 を参照しながら、さらに他の発明の実施の形態等について説明する。

【0385】

本発明は、例えば、仮想的なプライベートネットワークを構成するためのルー

タ装置、その通信方法などに関するものである。

【0386】

(従来の技術)

VPN (Virtual Private Network ; 仮想的プライベートネットワーク) を実現する従来技術は、おもに二通り存在する。

【0387】

一つは、プライベートなパケットを非プライベートなパケットでカプセル化する技術であり、もう一つは、プライベートなパケットを非プライベートネットワーク上のラベルスイッチパスにて転送する技術である。

【0388】

前者のカプセル化技術を用いたVPNでは、ラベルスイッチネットワークの存在を仮定せずともVPNが実現可能であるという柔軟性が利点である反面、ラベルスイッチネットワークにおけるデータ転送の高速性や通信品質の保証機能をVPN専用にご利用することができないという欠点があった。

【0389】

一方で、後者のラベルスイッチパス技術を用いたVPNでは、ちょうど前者における利点と欠点とが入れ替わることになる。

【0390】

(発明が解決しようとする課題)

ラベルスイッチネットワーク領域と非ラベルスイッチネットワーク領域とが混在している環境は、例えばインターネットサービスプロバイダーのバックボーンなどで、より広く一般的なものとなりつつある。

【0391】

本発明は、上記のような環境のもとで、一方でラベルスイッチが利用できるネットワーク領域では自動的にVPN専用のラベルスイッチパスを設定し、他方でラベルスイッチが利用できないネットワーク領域では自動的にカプセル化を施すようなVPN実現方法を与えることで、従来の二極的なVPN技術では得られなかった柔軟かつ高速で通信品質の保証をも期待できるVPNを提供可能なルータ装置、その通信方法を提供することを目的とする。

【0392】

(発明が解決しようとする課題)

本発明(請求項20)は、ラベルスイッチパスをその内部に設定可能なラベルスイッチネットワークと、それ以外の非ラベルスイッチネットワークとを含み、プライベートアドレス空間を共有する少なくとも二つのプライベートネットワークの間に位置する非プライベートネットワークにおいて、仮想的なプライベートネットワークを構成するための通信方法であって、前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータが、前記プライベートネットワークの一方から当該ルータへのラベルスイッチパスを設定するのに用いられた制御メッセージの内容に基づき、該ラベルスイッチパスから受信したパケットの宛先プライベートアドレスが前記プライベートネットワークの他方であることを認識し、受信した該パケットを前記プライベートネットワークの他方に対応する非プライベートアドレスでカプセル化して送信することを特徴とする。

【0393】

一方のプライベートネットワークからのラベルスイッチパスの設定に用いられる制御メッセージ中では、他方のプライベートネットワークを収容する境界ルータの非プライベートアドレスが指定され、どの仮想的なプライベートネットワーク(VPN)を構成するためのパスであるかが示されている。したがって、ラベルスイッチネットワークと非ラベルスイッチネットワークの境界にあるルータは、この制御メッセージの内容を記憶しておき、後にパケットを受信した時にこれを利用すれば、そのラベルスイッチパスから受信したパケット自体には宛先として未知のプライベートアドレスしか記入されていなくても、そのパケットは他方のプライベートネットワークを収容する境界ルータへ送るべきであると判断でき、この判断に基づいてパケットを非プライベートアドレスでカプセル化して送信することができる。これにより、非プライベートネットワークの全体がラベルスイッチネットワークではない(一方のプライベートネットワークはラベルスイッチネットワークに接続されているが他方のプライベートネットワークは非ラベルスイッチネットワークに接続されている)環境でも、一方のプライベートネット

ワークからのパケットを途中まではラベルスイッチパスを利用して他方のプライベートネットワークへ転送するVPNを実現することが可能になる。

【0394】

好ましくは、前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータが、前記プライベートネットワークの一方から当該ルータへのラベルスイッチパスを設定するのに用いられた制御メッセージの内容に基づき、自身が前記プライベートネットワークの境界ルータであると前記非プライベートネットワーク内に広告し、前記プライベートネットワークの他方を収容する前記非プライベートネットワークに接続されたルータは、この広告に基づいて前記プライベートネットワークの他方からのパケットを前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータの非プライベートアドレスでカプセル化して送信し、このカプセル化されたパケットを受信した前記ラベルスイッチネットワークと前記非ラベルスイッチネットワークとの境界に設置されたルータは、該カプセル化を解いて該パケットを当該ルータから前記プライベートネットワークの一方へのラベルスイッチパスへ送信するようにしてもよい。

【0395】

他方のプライベートネットワークを収容する境界ルータは、一方のプライベートネットワークを収容する境界ルータとして、一方のプライベートネットワークを直接収容しラベルスイッチネットワークに接続されたルータと、ラベルスイッチネットワークと非ラベルスイッチネットワークの境界にある広告を行ったルータ（プロキシ境界ルータと解釈することもできる）とを認識するが、ネットワーク的な距離の近さから、プロキシ境界ルータの方を選択して、カプセル化送信を行う。

【0396】

そして、ラベルスイッチネットワークと非ラベルスイッチネットワークの境界にあるルータは、受信したカプセル化パケットの送信元情報を調べることにより、このカプセル化パケットがどのVPNに属するプライベートネットワークを収容するどの境界ルータから送信されてきたかが分かる。この送信元情報に示され

る境界ルータが、一方のプライベートネットワークからラベルスイッチパスを介してパケットを受信した時にカプセル化送信の宛先とするルータであれば、そのラベルスイッチパスとちょうど逆方向のラベルスイッチパスであって、受信したカプセル化パケットの示すVPN用のものへ、カプセル化を解いて転送すれば良いことが分かる。

【0397】

このようにして、非ラベルスイッチネットワークに接続された他方のプライベートネットワークからのパケットも、途中からはラベルスイッチパスを利用してラベルスイッチネットワークに接続された一方のプライベートネットワークへ転送するVPNを実現することが可能になる。

【0398】

なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【0399】

また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【0400】

（発明の実施の形態）

図21は、ラベルスイッチネットワーク領域と非ラベルスイッチネットワーク領域との両方を含むネットワークの一例を示しており、かつ、このネットワーク上でVPNが構成されている様子を示している。

【0401】

一方のネットワーク領域2101は、ラベルスイッチが利用でき、かつVPN専用のラベルスイッチパスを設定することができるような、ネットワーク1901と同様の機能を持つネットワークである。

【0402】

他方のネットワーク領域 2102 は、前記ラベルスイッチを利用しないネットワークの一例である。

【0403】

両者を併せたものが、一つの非プライベートアドレス空間（以後、グローバルアドレス空間と呼ぶ）を共有する、非プライベートネットワークであるものとする。

【0404】

この非プライベートネットワークを挟んで、ネットワーク 2103 およびネットワーク 2104 が、一つのプライベートアドレス空間を共有する VPN を実現するものとする。

【0405】

ルータ装置 2105 は、非プライベートアドレス空間のラベルスイッチネットワーク 2101 とプライベートアドレス空間のネットワーク 2103 との境界に位置し、各々の側にネットワークインタフェースを備え、ラベルスイッチパス（以下、LSP と呼ぶ）の起点および終点となる機能を有するものである。

【0406】

ルータ装置 2106 は、非プライベートアドレス空間のラベルスイッチネットワーク 2101 と同アドレス空間の非ラベルスイッチネットワーク 2102 との境界に位置し、各々の側にネットワークインタフェースを備え、LSP の起点および終点となる機能を有し、かつ、パケットのカプセル化およびデカプセル化を行う機能を有するものである（以後、ラベルスイッチ境界ルータ装置と呼ぶ）。

【0407】

ルータ装置 2107 は、非プライベートアドレス空間の非ラベルスイッチネットワーク 2102 とプライベートネットワーク 2104 との境界に位置し、各々の側にネットワークインタフェースを備え、パケットカプセル化およびデカプセル化を行う機能を有するものである。

【0408】

LSP 2108 は、ルータ装置 2105 とルータ装置 2106 との間で、図 14 に示した処理手順と同様の処理手順にて設定された、VPN を構成するための

ものである。本 L S P には該 V P N のプライベートアドレスが付いたパケットが流れる。

【0409】

カプセル化パケット転送路 2109 は、同 V P N を構成するためのものである。

【0410】

ここでは、プライベートアドレスが付いたパケットが、非プライベートアドレスのパケットにてカプセル化されることにより、ルータ装置 2106 とルータ装置 2107 との間の非プライベートネットワーク上で、V P N のパケットが転送可能となっている。

【0411】

図 22 に、本発明によるラベルスイッチ境界ルータ装置について、図 21 に示した非プライベートネットワーク上のラベルスイッチ境界ルータ装置 2106 を例にとり、その一構成例を示す。

【0412】

I P 処理部 2201 は、ネットワーク層の処理を行うが、ここではグローバルアドレス空間に属するパケットを処理する。

【0413】

ルーティングテーブル 2202 は、グローバルアドレス空間に属する宛先アドレスや次段情報を管理する。

【0414】

L S P 制御部 2203 は、非プライベートネットワークを経由する L S P を設定するためのプロトコル（以後、L S P 設定プロトコルと呼ぶ）を実行する。

【0415】

ラベルスイッチネットワークインタフェース 2204 は、ラベルスイッチネットワーク 2101 に接続するためのネットワークインタフェースであり、グローバルアドレス空間に属するアドレスがつき、前記 L S P 設定プロトコルの通信インタフェースとしても機能する。

【0416】

本ネットワークインタフェース2204の下位レイヤとしては、LSPが設定可能なものであれば、どのようなものでも良い。例えばATM、フレームリレー、イーサネットなどが考えられる。

【0417】

ラベルスイッチネットワークインタフェースは、必要に応じて複数備えても良い。

【0418】

非ラベルスイッチネットワークインタフェース2205は、非ラベルスイッチネットワーク2102に接続するためのネットワークインタフェースであり、グローバルアドレス空間に属するアドレスがつく。

【0419】

本ネットワークインタフェース2205の下位レイヤとしては、LSPが設定可能である必要はなく、一般的にネットワーク層の下で利用されている任意の下位レイヤにて実施可能である。

【0420】

非ラベルスイッチネットワークインタフェースは、必要に応じて複数備えても良い。

【0421】

カプセル化部2206は、特定のLSPから受信したデータグラムを、非プライベートアドレス空間の packets ヘッダでカプセル化し、非ラベルスイッチネットワークインタフェース2205に転送可能とするためのものである。

【0422】

デカプセル化部2207は、非ラベルスイッチネットワークインタフェース2205にて受信した特定のデータグラムを、デカプセル化し（この場合、非プライベートアドレスが付いた packets ヘッダを除去し）、それらを特定のLSPに転送するためのものである。

【0423】

図23では、ラベルスイッチネットワークと非ラベルスイッチネットワークとが混在する非プライベートネットワーク上でVPNを実現するためのラベルスイ

ッチ境界ルータ装置の処理手順を、ラベルスイッチ境界ルータ装置 2106 の場合を例にとって説明する。

【0424】

ラベルスイッチ境界ルータ装置 2106 は、VPN メンバーシップ情報を獲得する（ステップ S2301）。かかる獲得の手段は、実施形態 2-1 と同様の VPN メンバーシップ配布手段で実現できる。これは VPN メンバーシップ情報データベースに保持する。

【0425】

図 21 において、例えば、ルータ装置 2105 に関する VPN メンバーシップ情報を得た場合、ルータ装置 2105 の VPN 識別子とグローバルアドレスとプライベートアドレスとを認識する。また、同様にルータ装置 2107 に関しても VPN メンバーシップ情報を得る。

【0426】

そこで、ラベルスイッチ境界ルータ装置 2106 は、自らも前記 VPN メンバーシップに参加宣言する（ステップ S2302）。

【0427】

この例では、自分と同じラベルスイッチネットワークの領域に属するルータ装置 2105 に関して、前記情報のうち、グローバルアドレスの項を自らのグローバルアドレスにて代入したものを、ルータ装置 2107 に通知する。通知の手段としては先述した VPN メンバーシップ配布手段を用いることができる。

【0428】

しかる後、ラベルスイッチ境界ルータ装置 2106 には、ラベルスイッチネットワーク 2102 を経由して、LSP 設定メッセージが到着する（ステップ S2303）。

【0429】

前記 LSP 設定メッセージの一例として、たとえばルータ装置 2105 によって送出されたものを考え、かつ、そのメッセージ中には、LSP のラベルに対応つける VPN フロー識別子において、当該 VPN の識別子ならびに LSP 終点エッジルータのグローバルアドレスとしてルータ装置 2107 のグローバルアドレ

スが記入されている場合を考える。

【0430】

このようなとき、ラベルスイッチ境界ルータ装置 2106 は、VPNメンバーシップ情報データベースを照合し前記設定メッセージを受け付けるかどうかを判断する（ステップ S2304）。

【0431】

VPNに関して前記設定メッセージを受け付けても良い場合の例としては、前記設定メッセージにて LSP のラベルに対応付けられている VPN フロー識別子について、VPN 識別子と LSP 終点エッジルータのグローバルアドレスとの組が VPN メンバーシップ情報データベースの中でマッチするエントリがあり、かつ、前記設定メッセージ中の起点エッジルータについても VPN メンバーシップ情報データベースの中でマッチするエントリがある、という条件が合致するときが挙げられる。

【0432】

設定メッセージを受け付けない場合は、終了する（ステップ S2305）。なお、LSP 設定プロトコルの仕様にも依存するが、もし N a c k を返す機構が定義されていれば、単に終了とせずに、N a c k を返しても良い。

【0433】

前記設定メッセージを受け付けた場合、LSP の設定完了通知を送出する（ステップ S2306）。

【0434】

この例では、LSP 2108 についてルータ装置 2105 からルータ装置 2106 への方向が設定完了したことに相当する。

【0435】

この LSP 上を流れてくるデータグラムは、パケットの宛先アドレスがプライベート空間アドレスとなっていて良い。

【0436】

ラベルスイッチ境界ルータ装置 2106 は、前記 LSP のラベルに対応付けられている VPN フロー識別子から、この LSP 上を流れてくるデータグラムをデ

ータグラムを転送すべきノードはルータ装置2107であるということを判断することができる。

【0437】

この判断に基づいて、ラベルスイッチ境界ルータ装置2106は、前記LSPから受けとったデータグラムを、非プライベートアドレスの packets ヘッダにてカプセル化し、転送を開始することができる（ステップS2307）。

【0438】

この例のカプセル化では、宛先アドレスはルータ装置2107のグローバルアドレスに、送信元アドレスはラベルスイッチ境界ルータ装置2106のグローバルアドレスに、それぞれ設定される。

【0439】

これは、カプセル化 packets 転送路2109についてルータ装置2106からルータ装置2107への方向が確立したことに相当する。

【0440】

ここまでの手順で、図21において、プライベートネットワーク2103からプライベートネットワーク2104への方向に、VPNのデータグラムを到達させることが可能になる。

【0441】

次に、ラベルスイッチ境界ルータ装置2106は、ルータ装置2105へ向けて、LSP設定メッセージを送出する（ステップS2308）。

【0442】

この例では、先述したLSP2108の逆方向の設定を要求する。

【0443】

しかる後、設定完了通知メッセージが到着すると（ステップS2309）、ラベルスイッチ境界ルータ装置2106は、ルータ装置2107から受信したカプセル化 packets をデカプセル化し、先述したルータ装置2105へ到達する当該LSPに転送を開始することができる（ステップS2310）。

【0444】

例えばルータ装置2107が一つのVPNだけを収容する場合は、ラベルスイ

タッチ境界ルータ装置2106は、カプセル化パケットの送信元アドレスとVPNメンバーシップ情報データベースとを調べることにより、そのパケットがどのVPNに属しているかを一意に定めることができ、かつ、デカプセル化してどのLSPに転送すべきかを一意に定めることができる。

【0445】

ただし、ルータ装置2107から受信するパケットは必ずしもカプセル化されたパケットではないが、そのときはデカプセル化や前記特定のLSPへの転送は行わない。

【0446】

また、カプセル化されたパケットであっても、もし、ルータ装置2107が複数の異なるVPNを収容している場合には、前記パケットがどのVPNに属しているかを定められるときに限り当該のLSPに転送するようにする。

【0447】

カプセル化パケットがどのVPNに属しているかを定める手段としては、例えば、パケットのカプセル時に付加的情報を挿入できる手段として、GREトンネル、IPsec、L2TPなどが適用できる。図21の例では、ルータ装置2107が前記付加的情報としてVPN識別子を挿入すれば良い。

【0448】

図24では、ラベルスイッチネットワークと非ラベルスイッチネットワークとが混在する非プライベートネットワーク上でVPNを実現するための、非ラベルスイッチネットワーク上のエッジルータ装置の処理手順を、図21のルータ装置2107の場合を例にとって説明する。

【0449】

カプセル化パケット転送路2109が設定されていない初期的状態において、ルータ装置2107は、先述したVPNメンバーシップ情報の収集手段を用いて、VPNメンバーシップ情報を新たに獲得する（ステップS2401）。

【0450】

図21の例では、ルータ装置2105およびルータ装置2106に関する情報を認識する。

【0451】

これらの情報エントリは、VPN識別子およびプライベートアドレスの項の値が同一であり、グローバルアドレスの項の値が異なっている。グローバルアドレスの項の値は、前者がルータ装置2105のグローバルアドレスであり、後者がルータ装置2106のグローバルアドレスとなっている。

【0452】

そこで、ルータ装置2017は、これらルータについて比較を行い、一方のルータを選択する（ステップS2402）。かかる選択の際はさまざまな基準を用いることが可能であり、例えばルータのグローバルアドレスについてネットワーク的な距離が近い方を選択するという基準が考えられる。

【0453】

ルータ装置2017は、プライベートネットワーク2104発のVPNのパケットを、非プライベートネットワーク2102を通過させるために、前記の選択したルータのグローバルアドレスにてカプセル化して、転送を開始する（ステップS2403）。

【0454】

前記のルータ選択の際、ネットワークの距離的により近いルータ2106を選択したのであれば、これは、図21におけるカプセル化パケット転送路2109についてルータ装置2107からルータ装置2106への方向が確立したことに相当する。

【0455】

ここまでの手順で、図21において、プライベートネットワーク2104からプライベートネットワーク2103への方向に、VPNのデータグラムを到達させることが可能になる。

【0456】

以上により、プライベートネットワーク2103とプライベートネットワーク2104との間で双方向のVPNが確立する。

【0457】

なお、以上の各機能は、ソフトウェアとしても実現可能である。

【0458】

また、各発明の各実施形態は、コンピュータに所定の手段を実行させるための（あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても実施することもできる。

【0459】

本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0460】

（発明の効果）

本発明によれば、一方のプライベートネットワークはラベルスイッチネットワークに接続されているが他方のプライベートネットワークは非ラベルスイッチネットワークに接続されている環境でも、ラベルスイッチネットワーク上ではラベルスイッチパスを利用したVPNを実現することが可能になる。

【0461】

【発明の効果】

本発明によれば、ロードバランスを実行する場合に、エッジルータの関与なしに、実際にマルチパス情報を持つルータ装置のみにて行うことができる。

【図面の簡単な説明】

【図1】

本発明の一実施形態に係るルータ装置を含むネットワークの構成例を示す図

【図2】

同実施形態に係るルータ装置の構成例を示す図

【図3】

カットスルーパス設定手順の一例を示すフローチャート

【図4】

カットスルーパス設定手順の他の例を示すフローチャート

【図5】

カットスルーパス設定手順のさらに他の例を示すフローチャート

【図 6】

ルーティングテーブルの一構成例を示す図

【図 7】

ラベルスイッチテーブルの一構成例を示す図

【図 8】

同実施形態に係るルータ装置の他の構成例を示す図

【図 9】

カットスループス再設定手順の一例を示すフローチャート

【図 10】

カットスループス再設定手順の他の例を示すフローチャート

【図 11】

ラベルスイッチテーブルの他の構成例を示す図

【図 12】

非プライベートネットワーク上に構築される VPN の一例を示す図

【図 13】

本発明の一実施形態に係る VPN エッジルータの構成例を示す図

【図 14】

エッジルータの処理手順の一例を示すフローチャート

【図 15】

VPN メンバーシップ情報の集合から成るデータベースの一例を示す図

【図 16】

グローバル処理モジュールが管理するラベルに関するテーブルの一例を示す図

【図 17】

バーチャルインタフェースの実装例を示す図

【図 18】

プライベート処理モジュールにおけるルーティングテーブルの一例を示す図

【図 19】

非プライベートネットワーク上に構築される複数の VPN の一例を示す図

【図 20】

本発明の一実施形態に係るVPNエッジルータの他の構成例を示す図

【図 21】

ラベルスイッチ領域と非ラベルスイッチ領域とを含む非プライベートネットワーク上に実現されるVPNの一例を示す図

【図 22】

本発明の一実施形態に係るラベルスイッチ境界ルータ装置の構成例を示す図

【図 23】

ラベルスイッチ境界ルータ装置の処理手順の一例を示すフローチャート

【図 24】

非ラベルスイッチルータ装置の処理手順の一例を示すフローチャート

【符号の説明】

101～104, 111～11N…ルータ装置

120…ネットワーク

201, 801…IP処理部

202, 802…カットスルー制御部

203, 803…ルーティングテーブル

204, 804…スイッチ部

211～21N, 221, 222, 811～81N, 821, 822…ネットワークインタフェース

231, 831…物理層処理部

232, 833…ラベル処理部

233, 834…ラベルテーブル

832…物理カウンタ

1221～122N, 1920～192N…エッジルータ装置

1211～121M, 1911～191M…ルータ装置

1201, 1901…非プライベートネットワーク

1231～123N, 1931～193N, 1991～199N…プライベートネットワーク

1251～125N, 1261～126N, 1305, 1314, 1950～

195N, 1971~197N, 1981~198N, 2005…ネットワーク
インタフェース

1301, 2001…グローバル処理モジュール

1302, 1312, 2002…IP処理部

1303, 1313, 2003…ルーティングテーブル

1304, 2004…LSP制御部

1306, 2006…スイッチ部

1311, 2011~201N…プライベート処理モジュール

1341~134N…バーチャルインタフェース

2101…ラベルスイッチネットワーク

2102…非ラベルスイッチネットワーク

2103, 2104…プライベートネットワーク

2105~2107…ルータ装置

2201…IP処理部

2202…ルーティングテーブル

2203…LSP制御部

2204…ラベルスイッチネットワークインタフェース

2205…非ラベルスイッチネットワークインタフェース

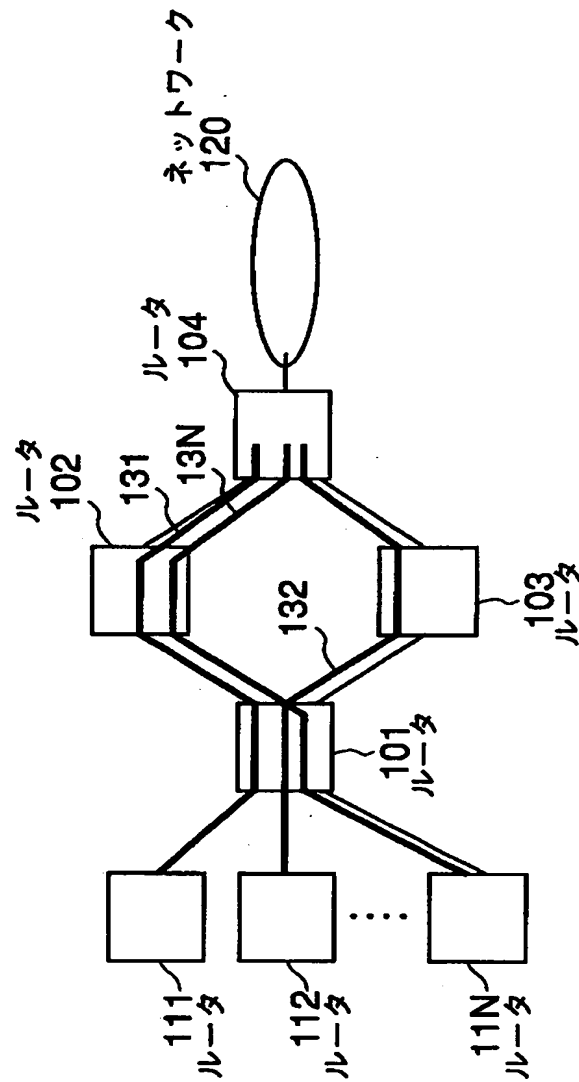
2206…カプセル化部

2207…デカプセル化部

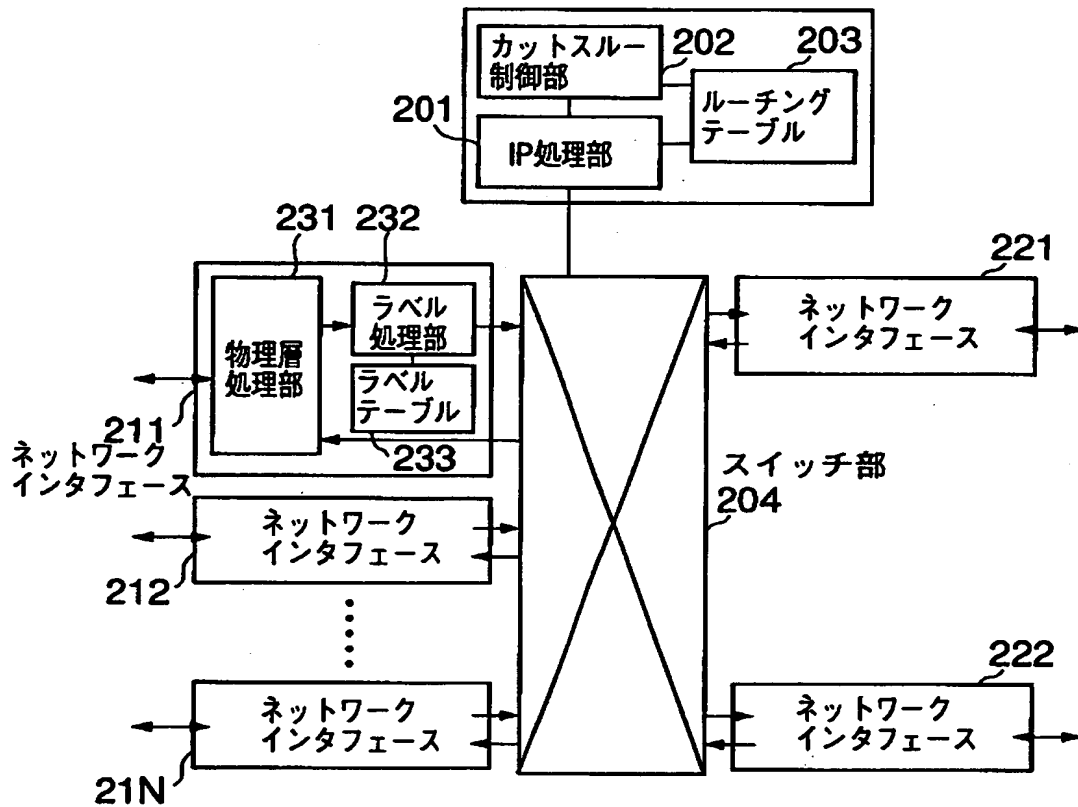
【書類名】

図面

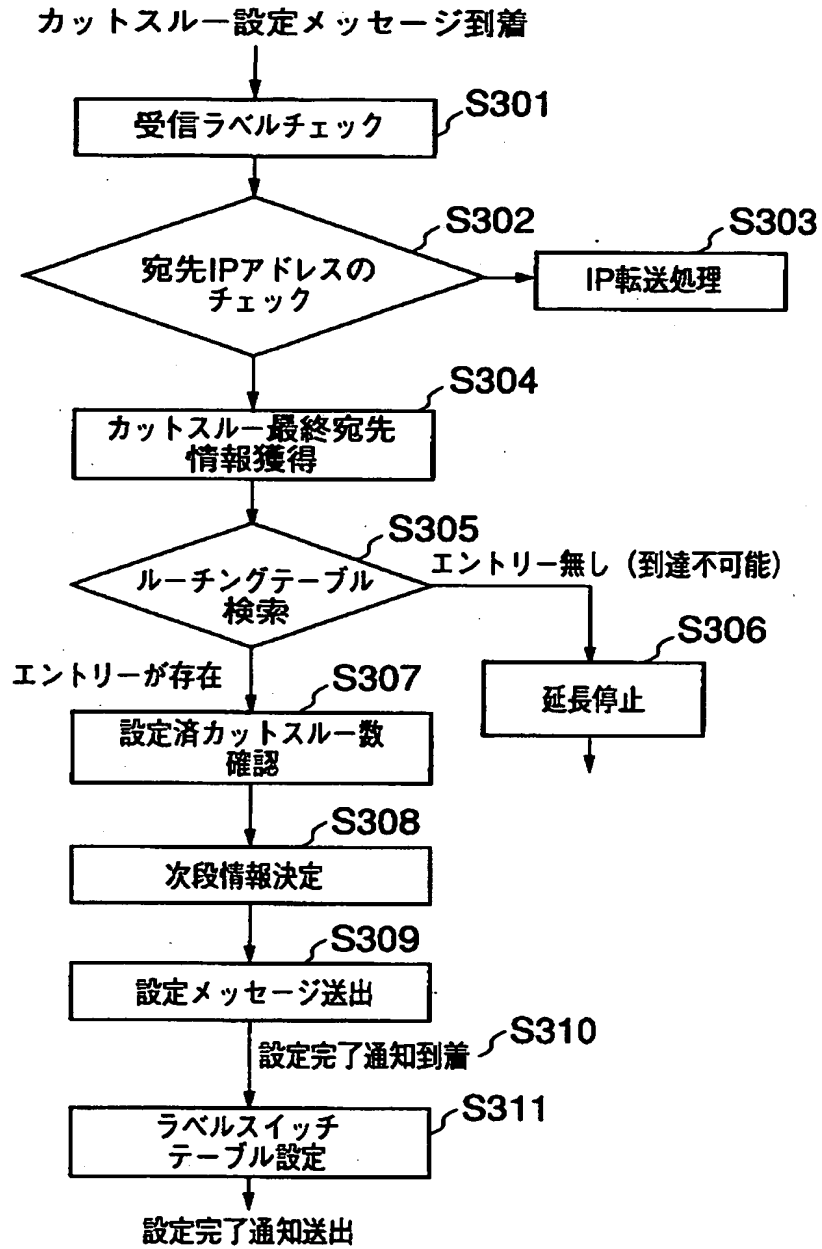
【図 1】



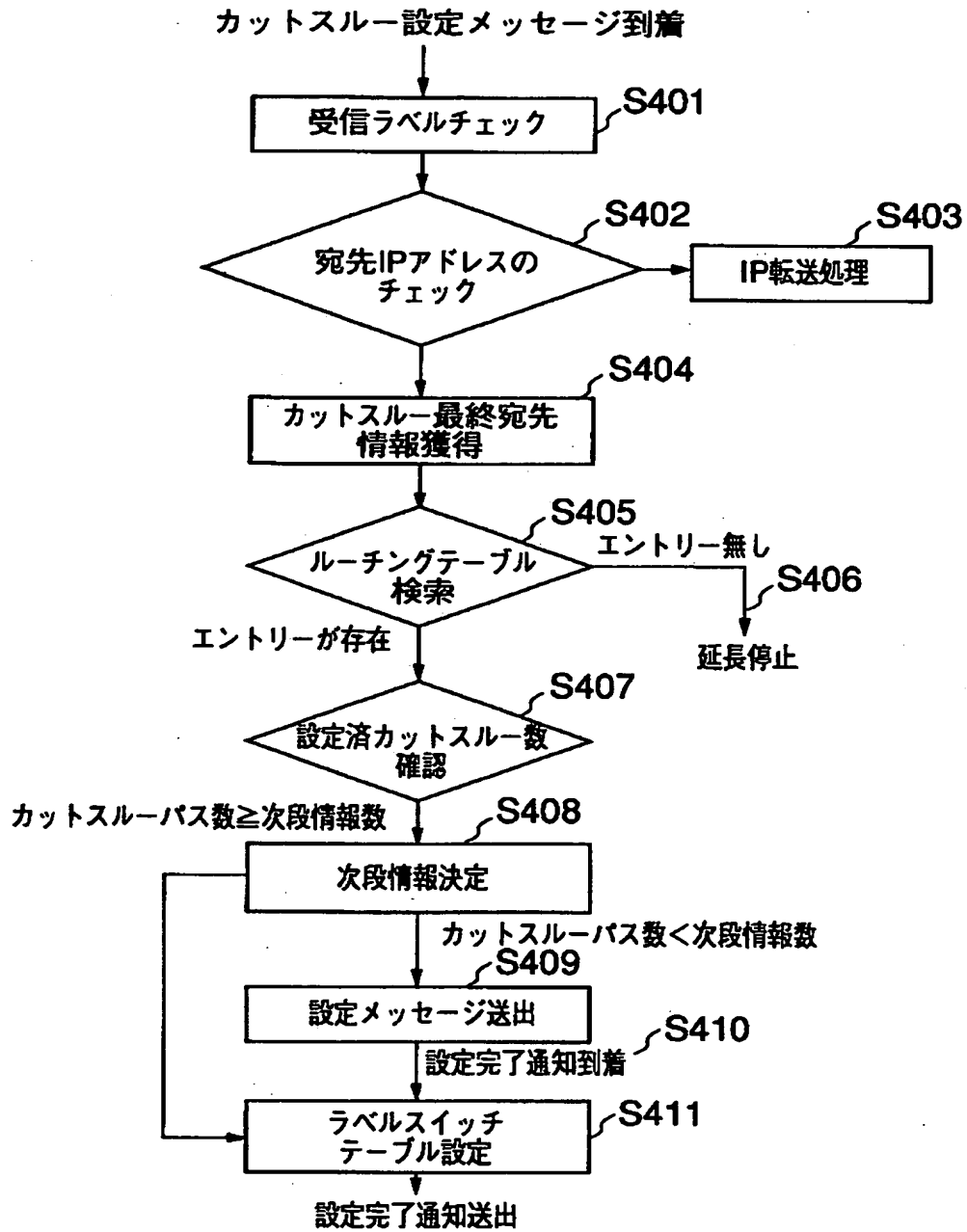
【図 2】



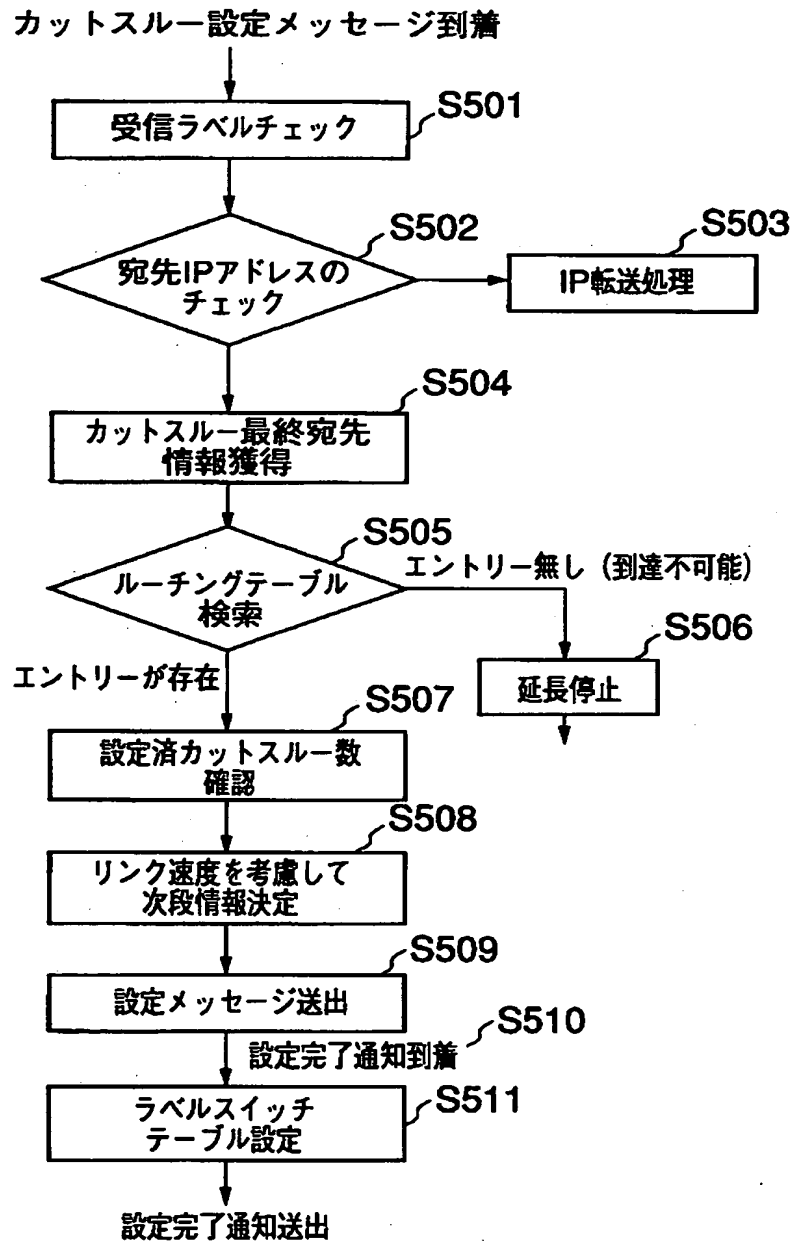
【図 3】



【図 4】



【図 5】



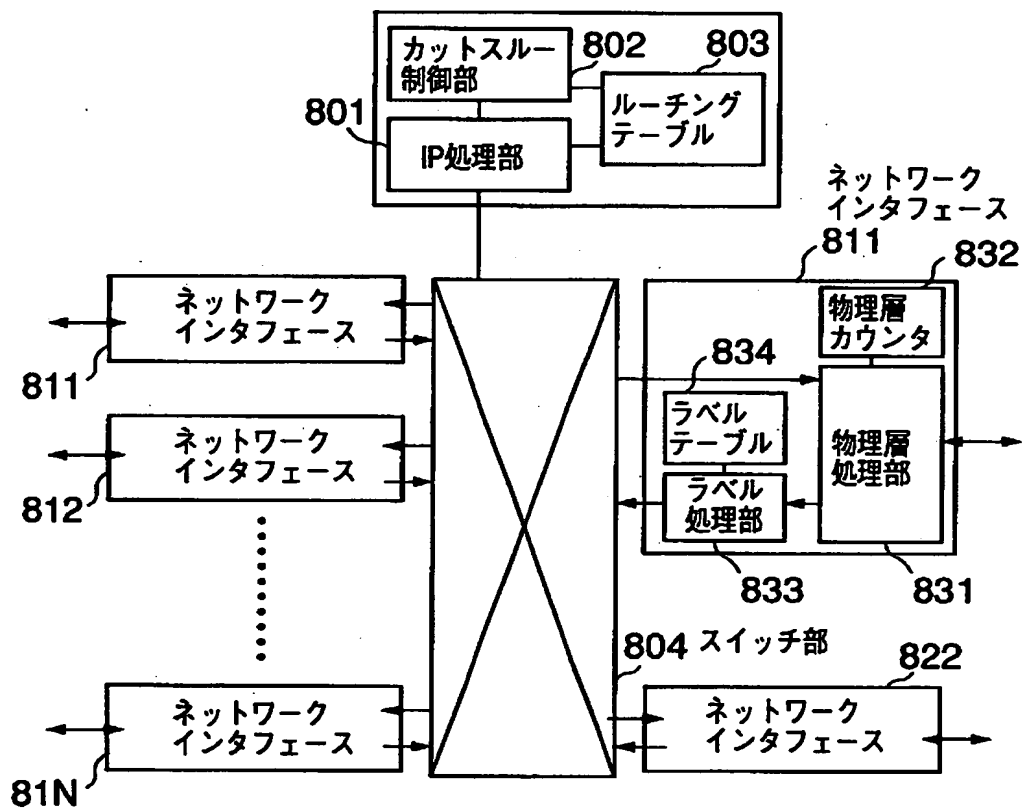
【図 6】

宛先アドレス	次段情報数	次段情報・・・	次段情報	カットスルー 設定数
⋮	⋮	⋮	⋮	

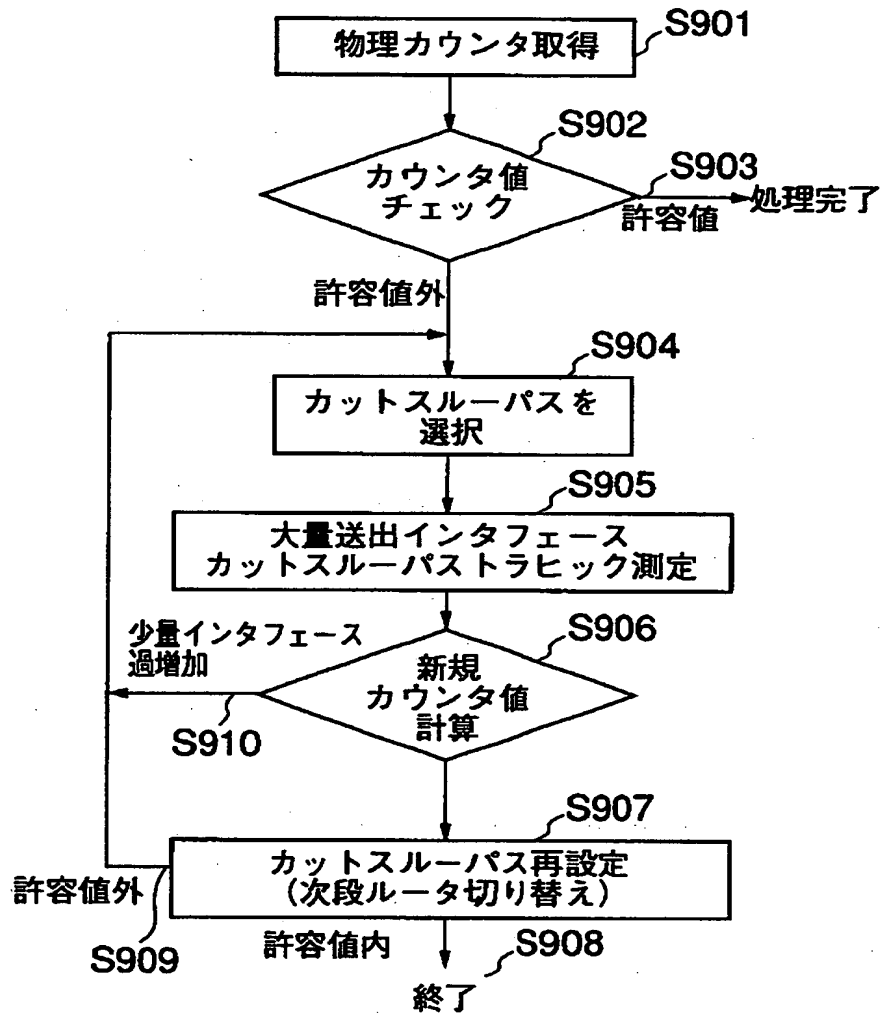
【図 7】

受信ラベル	出力インターフェース	送出ラベル値
⋮	⋮	⋮

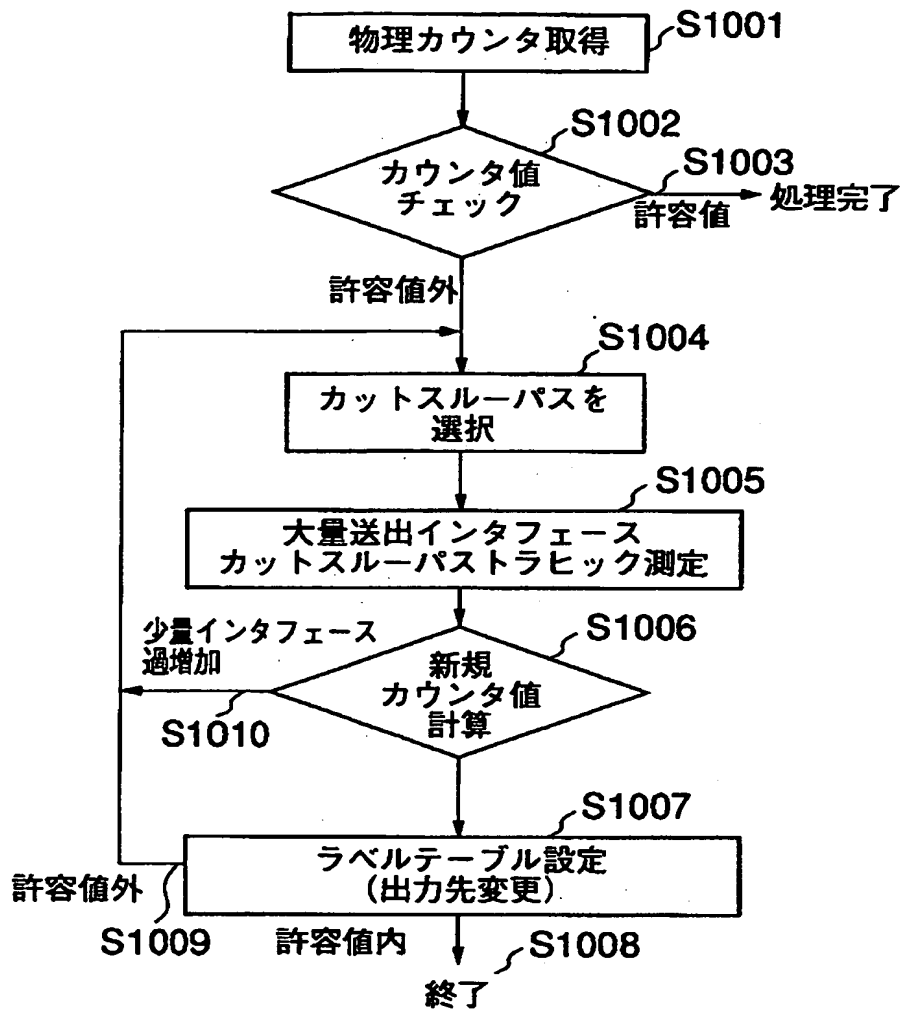
【図 8】



【図 9】



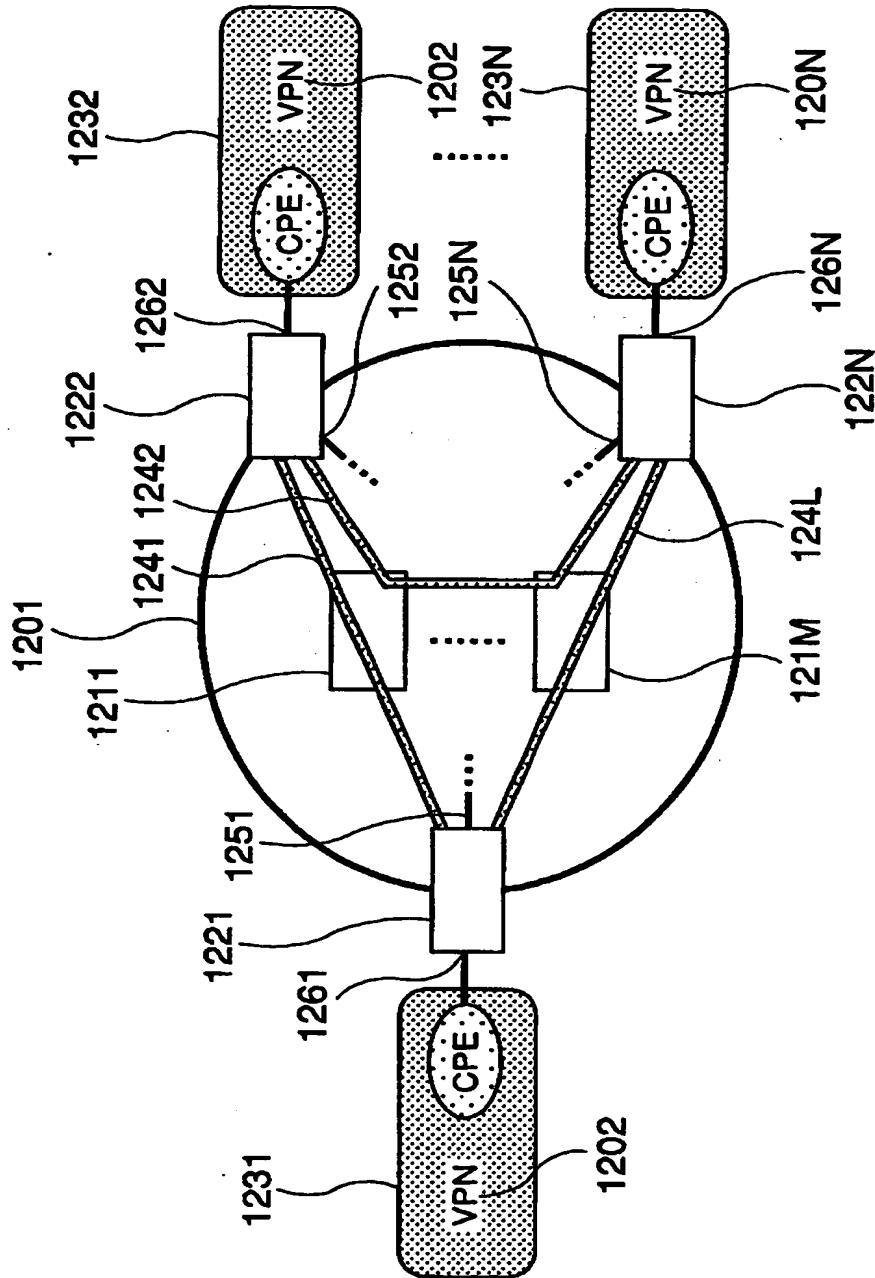
【図 10】



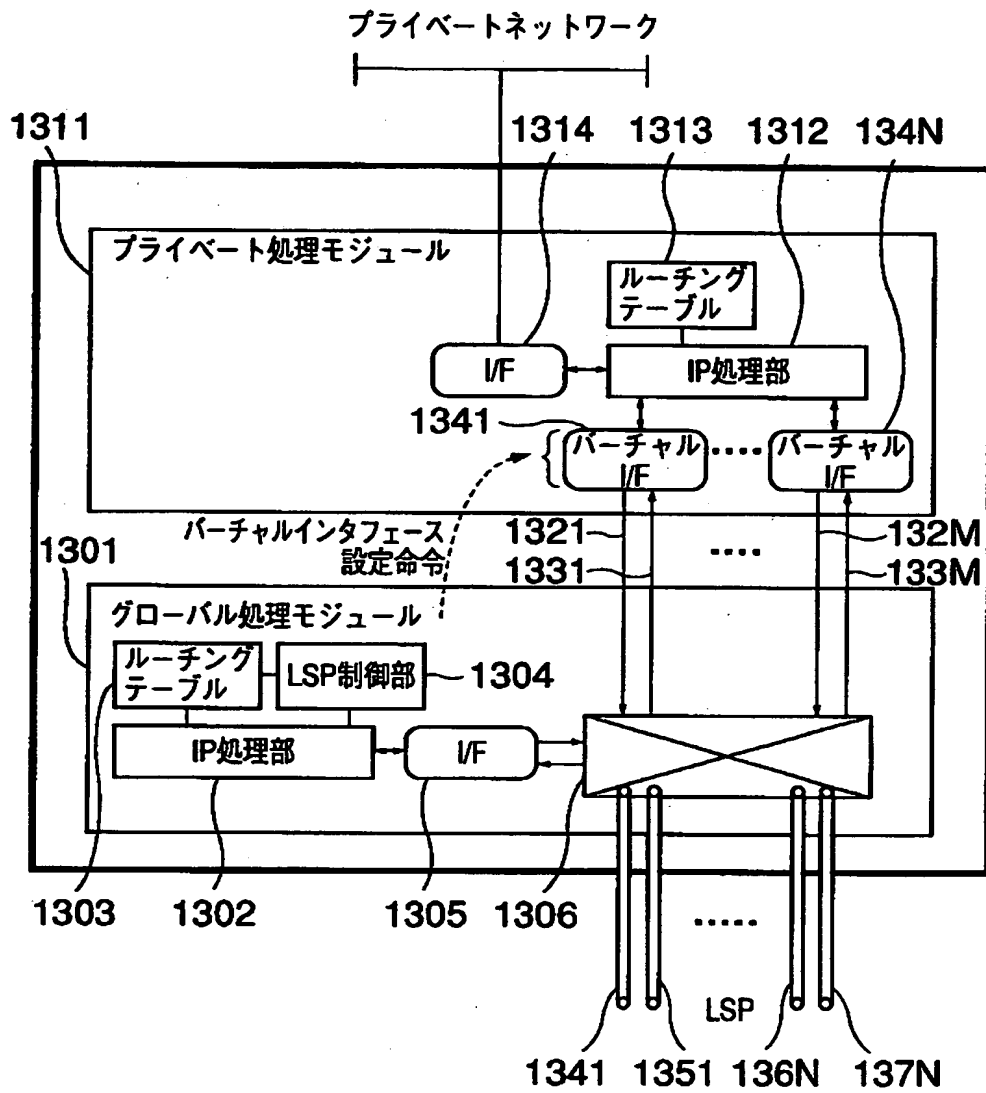
【図 11】

1101	1102	1103	1104
受信ラベル	出力インターフェース	送出ラベル値	受信フレーム数
⋮	⋮	⋮	

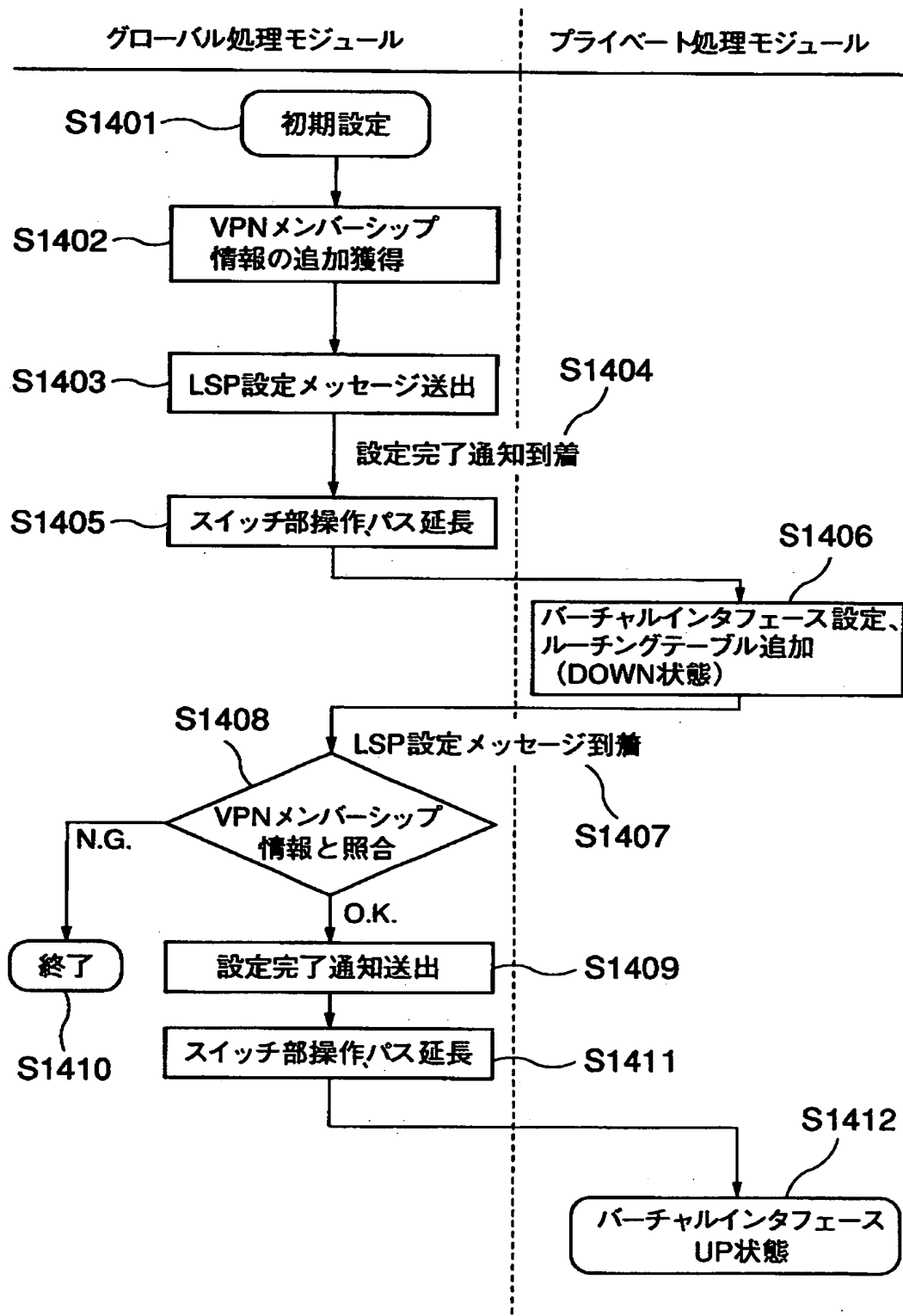
【図 12】



【図 13】



【図 14】



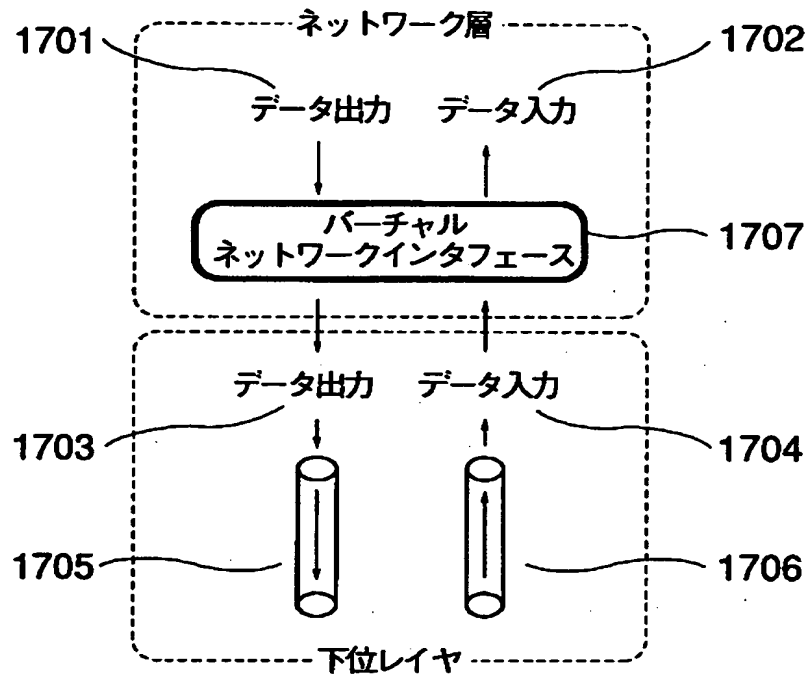
【図 15】

1501 VPN識別子	1502 プライベートアドレス	1503 グローバルアドレス
⋮	⋮	⋮

【図 16】

1601 VPNフロー識別子	1602 LSPのラベル値	1603 パスの識別子
⋮	⋮	⋮

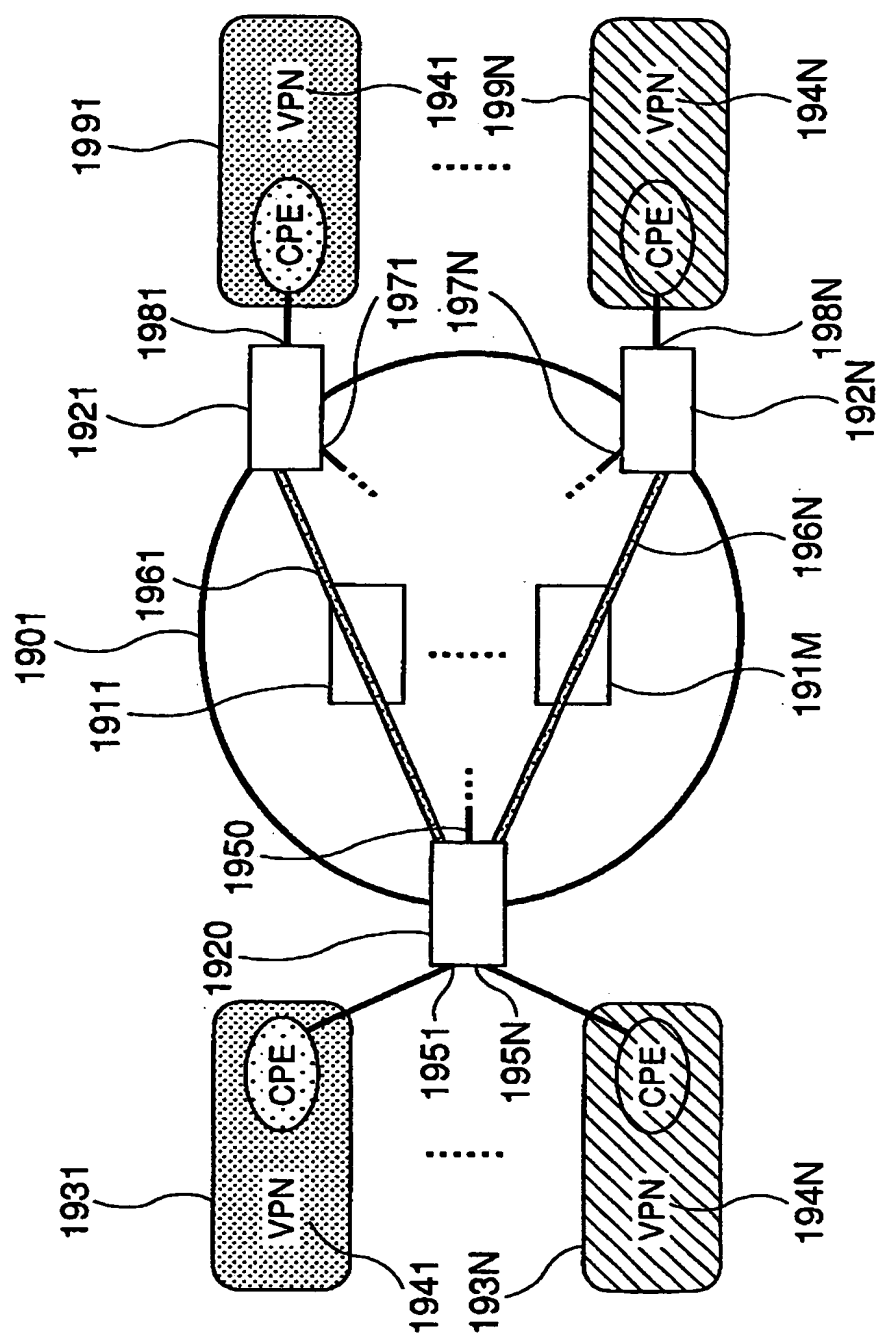
【図 17】



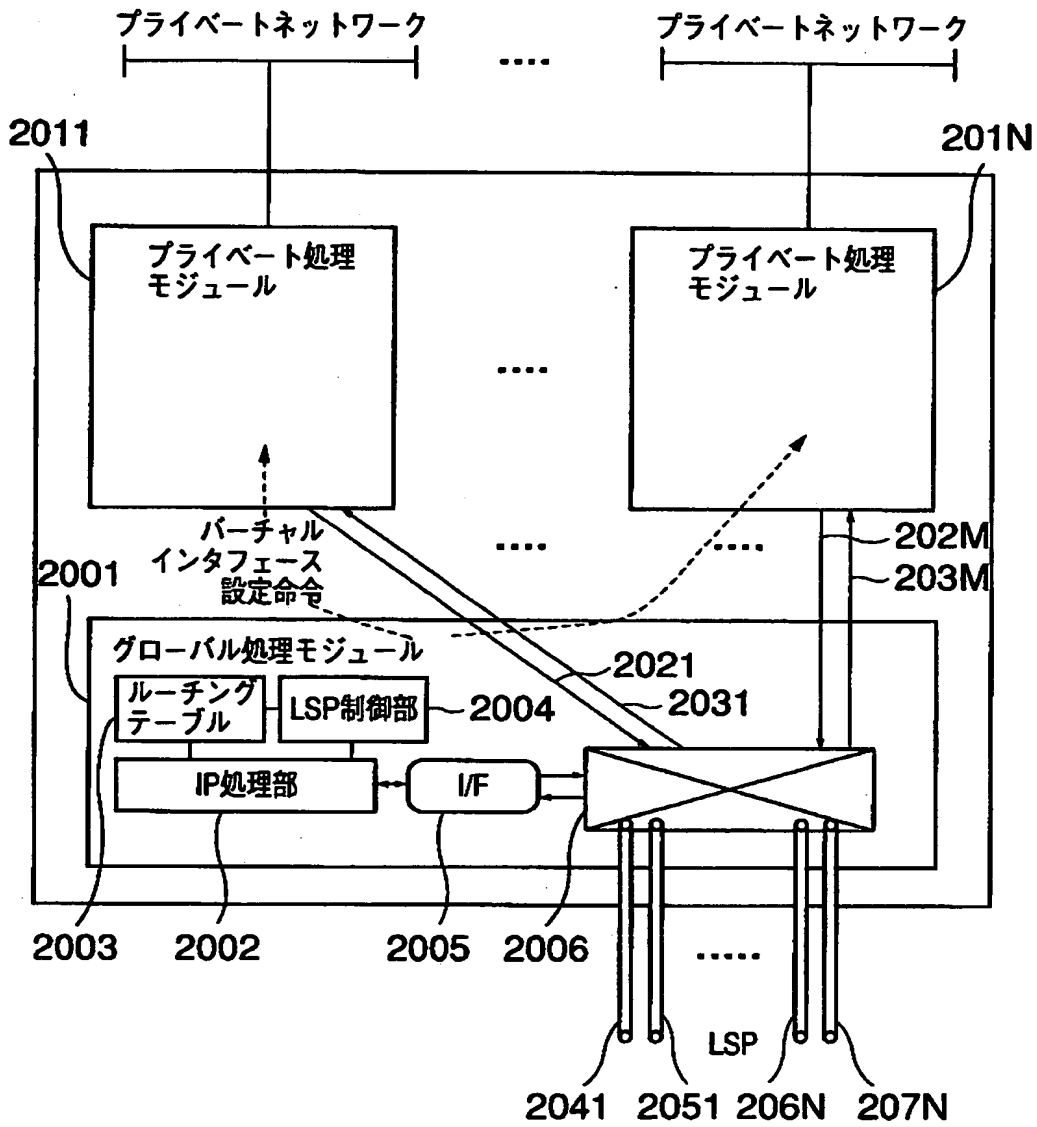
【図 18】

1801 宛先	1802 次段情報	1803 状態	1804 インタフェース名
⋮	⋮	⋮	⋮

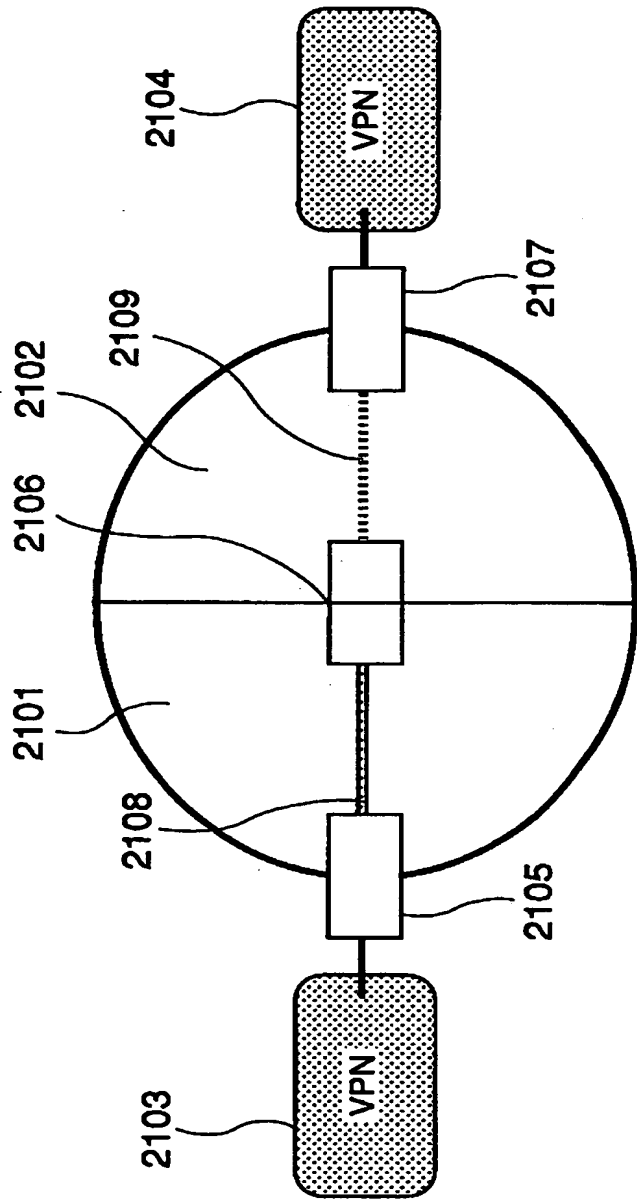
【図 19】



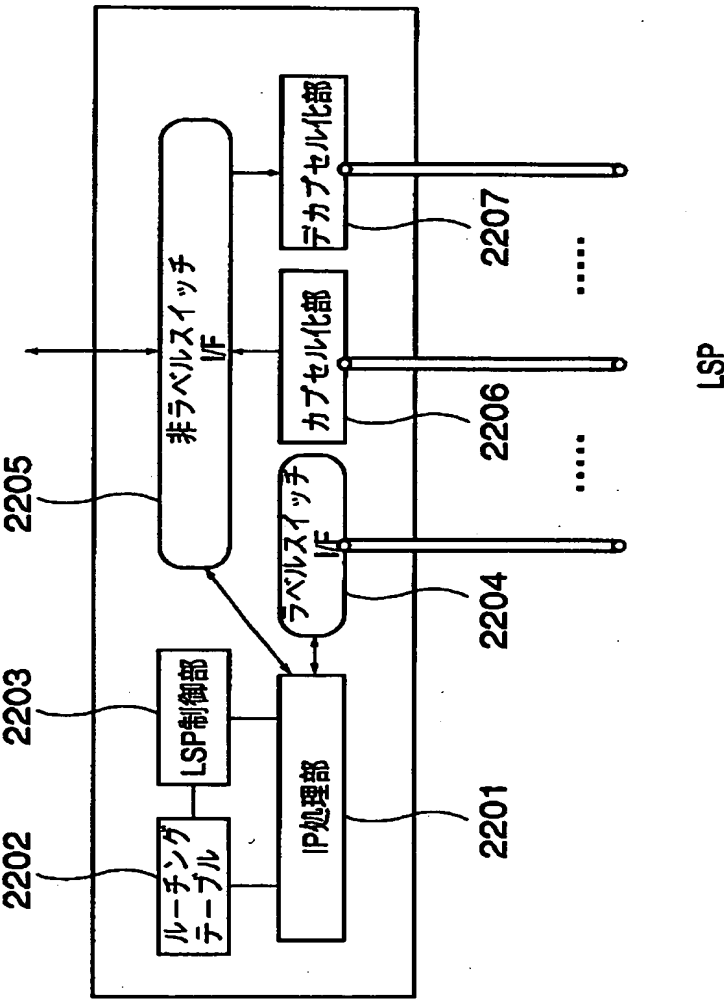
【図 20】



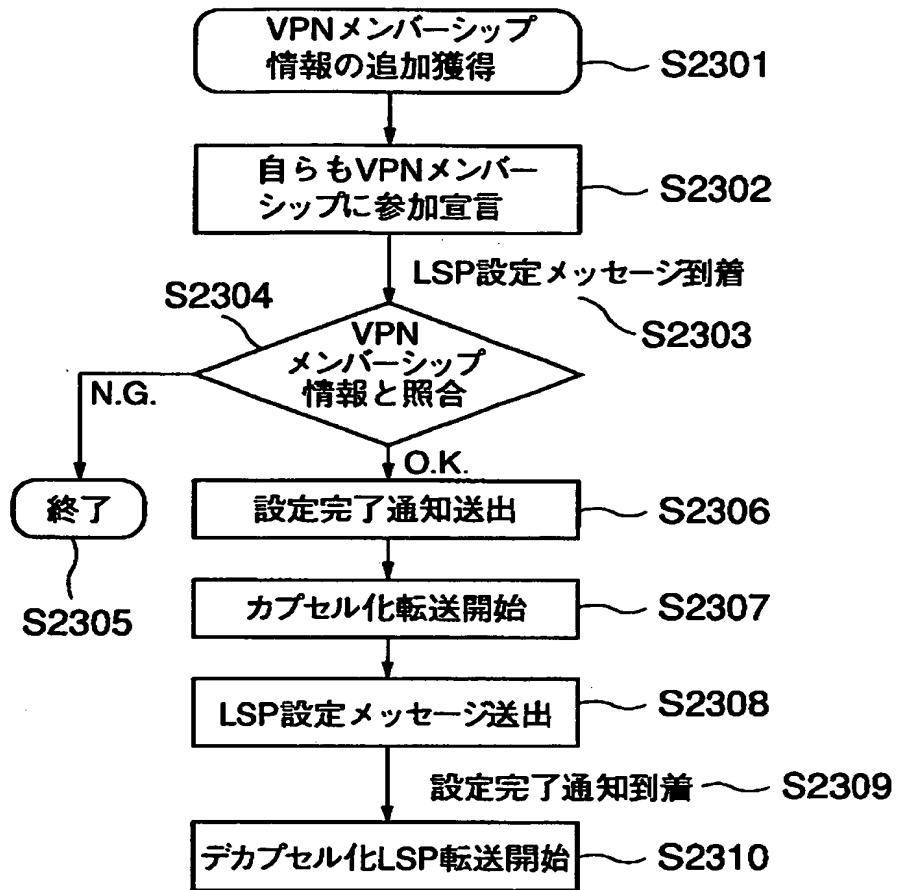
【図 21】



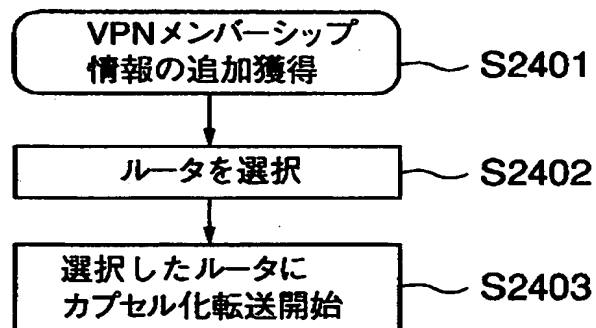
【図 22】



【図 23】



【図 24】



【書類名】 要約書

【要約】

【課題】 ロードバランスを実行する場合に実際にマルチパス情報を持つルータ装置にて行いエッジルータにて特別な処理が必要とならないようにしたカットスルーパス制御方法を提供すること。

【解決手段】 ある宛先のためのカットスルーパスを設定する際に、次段ルータとなり得るルータ装置が複数存在する場合、宛先を同じくする既設定のカットスルーパス数に基づいて、該次段ルータとなり得る各ルータ装置へのカットスルーパス数が均等的に配分されるように、該設定すべきカットスルーパスにおいて次段ルータとするルータ装置を選択する。

【選択図】 図 1

【書類名】 職権訂正データ
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000003078

【住所又は居所】 神奈川県川崎市幸区堀川町7番地

【氏名又は名称】 株式会社東芝

【代理人】 申請人

【識別番号】 100058479

【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮内外國
特許法律事務所内

【氏名又は名称】 鈴江 武彦

【選任した代理人】

【識別番号】 100084618

【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮内外國
特許法律事務所内

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮内外國
特許法律事務所内

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮内外國
特許法律事務所内

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮内外國
特許法律事務所内

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮内外國
特許法律事務所内

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】	100070437
【住所又は居所】	東京都千代田区霞が関3丁目7番2号 鈴榮内外國 特許法律事務所内
【氏名又は名称】	河井 将次

出 願 人 履 歴 情 報

識別番号 [000003078]

1. 変更年月日 1990年 8月22日
[変更理由] 新規登録
住 所 神奈川県川崎市幸区堀川町72番地
氏 名 株式会社東芝